

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein -- protein search, using sw model

Run on: April 13, 2006, 12:51:03 ; Search time 189 Seconds
(without alignments)
683.478 Million cell updates/sec

Title: US-10-686-490d-2
Perfect score: 1509
Sequence: 1 MAQVKANGITLEYEQGRRH.....ERHPLRIELIAGHAAAEA 294

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1509	100.0	294	8	ADO26338 Anti-Kasl
2	607.5	40.3	322	6	ADA34452 Acinetoba
3	597	39.6	299	8	ADR51257 Anti-biof
4	589.5	39.1	321	4	ABB09186 Acinetoba
5	586.5	38.9	312	4	ABB09185 Acinetoba
6	562.5	37.3	422	7	ABO73253 Pseudomon
7	483	32.0	360	8	ADL05923 M. catarr
8	314.5	20.8	267	2	AAW30522 Kurthia s
9	262.5	17.4	247	9	ABM95684 M. xanthu
10	227.5	15.1	271	7	ADI23937 Streptomy
11	220.5	14.6	276	6	ABU21130 Protein e
12	216.5	14.3	264	7	ADI23904 Streptomy
13	209.5	13.9	259	4	AAG64664 Hydroxyl
14	209.5	13.9	259	8	ADO70988 Stenotro
15	204	13.5	271	5	AAO22154 Ramoplan
16	203.5	13.5	284	9	ABE39939 L. pneumo
17	203.5	13.5	280	9	ABE36545 L. pneumo
18	198	13.1	232	5	AAE24160 Human 466
19	198	13.1	232	7	ADD12925 Human alp
20	198	13.1	232	8	ADT05828 Hidden Ma
21	198	13.1	232	8	ADT05777 Hidden Ma
22	196	13.0	283	2	AAR10975 Polymydr
23	192	12.7	287	7	ABO73720 Pseudomon
24	189.5	12.6	251	4	AAB60302 Alpha/bet

25	189	12.5	268	7	ABO83496 Pseudomon
26	185.5	12.3	233	7	ADB80253 Human 588
27	185	12.3	265	9	ABM92440 M. xanthu
28	184.5	12.2	325	3	AAG13760 Arabidops
29	184.5	12.2	647	5	ABB91674 Herbicida
30	184.5	12.2	647	8	ADN73561 Thale cre
31	183.5	12.2	251	4	ABN79868 Corynebac
32	183.5	12.2	251	4	AAG92372 C glucam
33	183.5	12.2	251	7	ADD13483 C. glutam
34	183	12.1	289	2	AAW99184 Rhodococc
35	182.5	12.1	271	9	ABM93099 M. xanthu
36	182	12.1	614	5	ABB91334 Herbicida
37	181.5	12.0	264	5	ABP40747 Staphyloc
38	181.5	12.0	264	8	ADS08054 Staphyloc
39	180.5	12.0	258	6	ABU42585 Protein e
40	180	11.9	338	7	ABO83828 Pseudomon
41	175.5	11.6	289	8	ADR13741 Amidase,
42	175.5	11.6	339	6	ABU21279 Protein e
43	175	11.6	261	9	ABM94555 M. xanthu
44	174.5	11.6	396	3	AAG48813 Arabidops
45	174.5	11.6	396	3	AAG48791 Arabidops

ALIGNMENTS

RESULT 1

ADO26338
ID ADO26338 standard; protein; 294 AA.

AC ADO26338;

DT 29-JUL-2004 (first entry)

DE Anti-Kaslauskas lipase protein.

ds; gene; enzyme; anti-Kaslauskas lipase; esterification; hydrolysis;
KW pharmaceutical; agricultural chemical; liquid crystal compound.

OS Synthetic.

PN EP1418237-A2.

PD 12-MAY-2004.

PF 06-OCT-2003; 2003EP-00022590.

PR 16-OCT-2002; 2002DE-01048166.

PA (FARB) BAYER CHEM AG.

PI Bosch B, Meissner R, Berendes F, Koch R;

DR WPI; 2004-378759/36.

DR N-PSDB; ADO26337.

XX New nucleic acid encoding anti-Kaslauskas lipase and derived enzymes,
PT useful for stereospecific hydrolysis and synthesis of aralkyl esters,
intermediates for pharmaceuticals and plant protection agents.

XX Claim 11; Page 16-17; 22pp; German.

PS The present invention provides the protein and coding sequences of an
anti-Kaslauskas lipase. These are used as catalysts in esterification or
CC hydrolysis reactions for preparation of enantiomeric aralkanols, or their
CC esters, useful for preparing pharmaceuticals or agricultural chemicals,
CC also liquid crystal compounds. The present sequence is the protein of the
invention.

XX Sequence 294 AA;

Query Match 100.0%; Score 1509; DB 8; Length 294;
Best Local Similarity 100.0%; Pred. No. 1e-151;

Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVKANGITLVEBOGHRHPSMLIMGLGGQLIDWPPEFIRGLAERGFVPCFNRDA 60
 DB 1 MAQVKANGITLVEBOGHRHPSMLIMGLGGQLIDWPPEFIRGLAERGFVPCFNRDA 60
 QY 61 GLSTKLEGVKKPNIARVFLASMLKPRVPYTTDDMALDTVGLMDALGIESHTVVGVSMG 120
 DB 61 GLSTKLEGVKKPNIARVFLASMLKPRVPYTTDDMALDTVGLMDALGIESHTVVGVSMG 120
 QY 121 GMIAQILGAKHGERVKSITLMTSSGNRMPAPRQVLOKFMKRVKPSMDKEEWIKYNLEL 180
 DB 121 GMIAQILGAKHGERVKSITLMTSSGNRMPAPRQVLOKFMKRVKPSMDKEEWIKYNLEL 180
 QY 181 LTTIGSPGLDREKALDVKRSIERCLCPEGTORQALAILQSGSRVKLLRRIAVPTLVISG 240
 DB 181 LTTIGSPGLDREKALDVKRSIERCLCPEGTORQALAILQSGSRVKLLRRIAVPTLVISG 240
 QY 241 AEDPLLPYOCGRDIADHIFGAFELIEGMDHIDPERHIPRLIELIAGHAAAAA 294
 DB 241 AEDPLLPYOCGRDIADHIFGAFELIEGMDHIDPERHIPRLIELIAGHAAAAA 294

RESULT 2
 ADA34452
 ID ADA34452 standard; protein; 322 AA.
 AC ADA34452;
 DT 20-NOV-2003 (first entry)
 DE Acinetobacter baumannii protein #1613.
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.
 OS Acinetobacter baumannii.
 PN US6562958-B1.
 XX US6562958-B1.
 PD 13-MAY-2003.
 PF 04-JUN-1999; 99US-00328352.
 PR 09-JUN-1998; 98US-0088701P.
 PA (GENO-) GENOME THERAPEUTICS CORP.
 PI Breton G, Bush D;
 DR WPI; 2003-576092/54.
 DR N-PSDB; ADA30326.

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, or as biocontrol agents for plants.
 Example; SEQ ID NO 5739; 328pp; English.
 The invention relates to isolated Acinetobacter baumannii nucleic acids. The A. baumannii nucleic acids and polypeptides are useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, to detect the presence of A. baumannii and other Acinetobacter species in a sample, in screening compounds for the ability to interfere with the A. baumannii life cycle or to inhibit A. baumannii infection, and as biocontrol agents for plants. The present sequence represents the amino acid sequence of an A. baumannii protein.
 Sequence 322 AA;
 Query Match 40.3%; Score 607.5; DB 6; Length 322;

Best Local Similarity 46.9%; Pred. No. 1.4e-55;
 Matches 134; Conservative 49; Mismatches 98; Indels 5; Gaps 4;

QY 7 NGITLVEBOGHRHPSMLIMGLGGQLIDWPPEFIRGLAERGFVPCFNRDAGLSTKL 66
 DB 33 NGIELHVEVGKPEHTITLLIMGLGAQMLFPDFFCKSLIDQGFVRVIRPDNRDGLSSKV 92
 QY 67 --EGVKKPNIARVFLASMLKPR--VPYTTDDMALDTVGLMDALGIESHTVVGVSMGMI 123
 DB 93 RHQG--KRLLTWKIMGRFALGURNQCAPYTLVDMADDVSMLLDRLGVSXKAVLGASMGMI 151
 QY 124 AQILGAKHGERVKSITLMTSSGNRMPAPRQVLOKFMKRVKPSMDKEEWIKYNLELTT 183
 DB 152 AQILAAKYPEKVEKGLMFTNNQPPFLPPPKQLLSLIGKPESDRDESGIVNHSKLPL 211
 QY 184 IGSPG--LDREKALDVKRSIERCLCPEGTORQALAILQSGSRVKLLRRIAVPTLVISGAE 242
 DB 212 IGSPGYINHIEAVOTARKLYQRSYYPAGVLOQFLAILCTGSLQLDREIKQPTLVLHGSR 271
 QY 243 DPLLPYOCGRDIADHIFGAFELIEGMDHIDPERHIPRLIELIAGH 288
 DB 272 DRLLPPSHGKAVAKAISGAFELIDGMGHDIPAHFIPQLSGLFAHH 317

RESULT 3
 ADR51257
 ID ADR51257 standard; protein; 299 AA.
 AC ADR51257;
 DT 04-NOV-2004 (first entry)
 DE Anti-biofilm polypeptide #24.
 KW antibacterial; antiinflammatory; anti-biofilm polypeptide;
 KW humoral response; environmental sample; water; liquid; soil; air;
 KW biological sample; medical device; pharmaceutical; food product;
 KW cosmetic; hygiene product; water treatment device; water transport; pulp;
 KW paper processing; paper recycling equipment; toothpaste; chewing gum;
 KW mouthwash; dental cleaner; caries; plaque; gingivitis; periodontitis;
 KW oropharyngeal colonization; gastric colonization; thermostability;
 KW thermotolerance.
 OS Unidentified.
 XX Unidentified.
 FH Key Location/Qualifiers
 FT Domain 49..288
 FT /note= "alpha/beta hydrolase fold domain"
 PN WO2004066945-A2.
 PD 12-AUG-2004.
 PF 26-JAN-2004; 2004WO-US002242.
 XX 24-JAN-2003; 2003US-0442794P.
 PR (DIVE-) DIVERSA CORP.
 PA Barton N, Robertson D, Chang K, Elkins J;
 XX WPI; 2004-625302/60.
 DR N-PSDB; ADR51256.
 XX New anti-biofilm enzymes such as esterase, glycosidase, amylase, useful for controlling biofilms on a wide range of household, industrial, medical surfaces.
 PS Claim 45; SEQ ID NO 48; 413pp; English.
 CC The invention relates to an isolated or recombinant anti-biofilm polypeptide (I) having at least 50-100% sequence identity to 58 fully defined sequences (S1), over region of at least 100 residues, where the

sequence identities is determined by analysis with a sequence comparison algorithm or by visual inspection. (I) is useful for making an antibody by administering to a non-human animal (I) or a DNA (II) encoding (I), to generate a humoral response. Oligonucleotide primers (III) are useful for isolating or recovering (II) encoding (I) from an environmental sample comprising water, liquid, soil, air or biological sample derived from a bacterial, protozoan, insect, yeast, plant, fungal or mammalian cell. (I) is useful for making a small molecule by providing several biosynthetic (I) capable of synthesizing or modifying a small molecule, and reacting the substrate with enzymes. (I) is useful for preventing the growth of a biofilm on a medical device, pharmaceutical, food product, device for making a food, cosmetic, hygiene product, water treatment device, water transport or storage device or pulp and paper processing and paper recycling equipment. The pharmaceutical is a tablet, pill, implant, suppository, inhaler, spray or ointment. The biofilm is from *Pseudomonas* or *Staphylococcus*, which involves administering to the biofilm (I) having esterase activity. (I) is useful in cleaning and decontaminating hard surfaces such as floors, equipment, machinery and industrial water treatments. (I) is useful as preservatives in food, medicinal, hygiene, and cosmetic products, useful in toothpaste, chewing gum, mouthwashes, dental cleaners. (I) is useful for controlling dental biofilms associated with caries, plaque, gingivitis, and periodontitis. (I) is useful for treating oropharyngeal and gastric colonization by pathogenic microorganisms. The activity of (I) is thermostable or thermotolerant. This sequence corresponds to a polypeptide of the invention.

XX Sequence 299 AA;

Query Match 39.6%; Score 597; DB 8; Length 299;

Best Local Similarity 45.1%; Pred. No. 1.7e-54;

Matches 128; Conservative 43; Mismatches 111; Indels 2; Gaps 1;

QY 9 ITLEYEQHRRHPSMLLMGLGQQLIDWPEEFIRGLAERGFVFCFDRDAGLSTKLEG 68
DB 8 IREYETFGHPDDPAIVLMGLGQQLIDWPEAFCRMADAGHYVVRFDNRDGLSTHLDH 67
QY 69 VKPNRIARVFLASMGKRPVYTLDDMALDTVGLMDALGISTHVVVSGMGMTAIIIG 128
DB 68 LPRPNLPLAALROALRPVRSYTLDDMDADDVAGLDLALNIIQAHVVGSMGMAQLLA 127
QY 129 AKHGERVKSITLMTTSSGNPRMPAPRPOVLQKFMVRPKSMDEWIKYNLELLTTIGSPG 188
DB 128 ARHATKVRSLTLLMTSSGNPRMPAPRPOVLQKFMVRPKSMDEWIKYNLELLTTIGSPG 187
QY 189 LDREKIALD--VRKSIERCLCEGTQORQALAIQSGSRVKKLRIIAVPTLVISGAEPL 246
DB 188 YPKPEALRIIVAEGRDFRAPHFAGFWRQLHVAIAPSRAPLLPRIKQPADVIHGADLIV 247
QY 247 PYOCGRDIADHIPGAFELIEGMGHDIPERHPIRLIELIAGHAA 290
DB 248 PVAARDLVRRLPNATLIDVPGMGHDFPTEIMPRIARRIVETAA 291

RESULT 4
ID ABB09186
XX ABB09186 standard; protein; 321 AA.
AC ABB09186;
XX ABB09186;
XX 04-JUL-2002 (first entry)
XX Acinetobacter calcoaceticus carboxyl esterase #2.
DE Acinetobacter calcoaceticus carboxyl esterase #2.
KW *Pseudomonas aeruginosa*; esterase; estA; enzyme; enantiomeric;
KW optical active carboxylic acid; asymmetric hydrolysis; hypotensive;
KW carboxylic ester racemate; hypertension; analapril; captopril.
XX Acinetobacter calcoaceticus.
OS Acinetobacter calcoaceticus.
XX WO200132847-A1.
XX 10-MAY-2001.
PD
XX

PF 01-NOV-2000; 2000WO-KR001243.
XX
PR 01-NOV-1999; 99KR-00047927.
XX
PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.
PA (BIOH-) BIOHOLDINGS INC.
XX
XX Kim C, Rhee S, Song K, Lee J, Boyapaty G;
PI WPI; 2001-316432/33.
XX
DR Novel esterase derived from *Pseudomonas aeruginosa* useful for producing
PT optical active carboxylic acids and pharmaceuticals, preferably
PT hypertension treatments drugs like analapril or captopril.
XX
XX Example 6; Fig 3; 40pp; English.
PS
XX The present invention describes an esterase (I) derived from *Pseudomonas*
CC *aeruginosa* which has a sequence comprising 315 amino acids, and produces
CC optical active carboxylic acids and enantiomeric isomers by the
CC asymmetric hydrolysis of the carboxylic ester racemates. (I) has
CC hypotensive activity. (I) is useful for producing optical active
CC carboxylic acids, preferably optical active (R)-carboxylic acids from
CC carboxylic acid ester racemates. (I) is useful for the production of
CC various kinds of physiologically active pharmaceuticals, specially
CC hypertension treatments drugs like analapril or captopril. (I) has
CC excellent ability for producing optically active carboxylic acids from
CC carboxylic acid esters. Production of optically active carboxylic acids
CC by (I) is highly more selective, simple and environmentally favourable
CC when compared to conventional techniques. (I) is thermally stable at high
CC temperatures above 70 degrees Celsius. The present sequence represents a
CC carboxyl esterase amino acid sequence given in comparison with the
CC esterase of the present invention
XX
XX Sequence 321 AA;

Query Match 39.1%; Score 589.5; DB 4; Length 321;

Best Local Similarity 45.5%; Pred. No. 1.2e-53;

Matches 130; Conservative 53; Mismatches 100; Indels 3; Gaps 3;

QY 6 ANGTLEYEQHRRHPSMLLMGLGQQLIDWPEEFIRGLAERGFVFCFDRDAGLSTK 65
DB 33 SNGLEHVEVGGNPDHPHILLMGLGQQLIDWPEEFIRGLAERGFVFCFDRDAGLSTK 92
QY 66 LEGV-KKPNRIARVFLASMGK-KRPVYTLDDMALDTVGLMDALGISTHVVVSGMGMI 123
DB 93 IRRHKGRLNTLKLMSRFTLGLNGOAPYTLYDMAEDVSLLEAMRIKKYNVIGASMGMI 152
QY 124 AQILGAKHGERVKSITLMTTSSGNPRMPAPRPOVLQKFMVRPKSMDEWIKYNLELLTT 183
DB 153 AQITAAKYPEKVEKALMFTSNNQPLPPPPPKQLFSLIGKPKSDEDOGINHSLKLF 212
QY 184 IGSPGLDREKIALD--VRKSIERCLCEGTQORQALAIQSGSRVKKLRIIAVPTLVISGAE 242
DB 213 IGSPGVNQVEAIQTARKLYQSRVHPAGVLOQFLAILCTGSLQLDQISQTLVHGR 272
QY 243 DPLLPYOCGRDIADHIPGAFELIEGMGHDIPERHPIRLIELIAGH 288
DB 273 DRLLPPSHGKAVAKAIGAKFELIQGMGHDIPPHFIPQLSYLFAHH 318

RESULT 5
ID ABB09185
XX ABB09185 standard; protein; 312 AA.
AC ABB09185;
XX ABB09185;
XX 04-JUL-2002 (first entry)
XX Acinetobacter calcoaceticus carboxyl esterase #1.
DE Acinetobacter calcoaceticus carboxyl esterase #1.
KW *Pseudomonas aeruginosa*; esterase; estA; enzyme; enantiomeric;
KW optical active carboxylic acid; asymmetric hydrolysis; hypotensive;

KW carboxylic ester racemate; hypertension; analapril; captopril.

XX Acinetobacter calcoaceticus.

XX WO200132847-A1.

XX 10-MAY-2001.

XX 01-NOV-2000; 2000WO-KR001243.

XX 01-NOV-1999; 99KR-00047927.

XX (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.

XX (BIOH-) BIOHOLDINGS INC.

XX Kim C, Rhee S, Song K, Lee J, Boyapaty G;

XX WPI; 2001-316432/33.

XX Novel esterase derived from *Pseudomonas aeruginosa* useful for producing
PT optical active carboxylic acids and pharmaceuticals, preferably
PT hypertension treatments drugs like analapril or captopril.
XX
XX Example 6; Fig 3; 40pp; English.

XX The present invention describes an esterase (I) derived from *Pseudomonas*
XX *aeruginosa* which has a sequence comprising 315 amino acids, and produces
CC optical active carboxylic acids and enantiomeric isomers by the
CC asymmetric hydrolysis of the carboxylic ester racemates. (I) has
CC hypotensive activity. (I) is useful for producing optical active
CC carboxylic acids, preferably optical active (R)-carboxylic acids from
CC carboxylic acid ester racemates. (I) is useful for the production of
CC various kinds of physiologically active pharmaceuticals, specially
CC hypertension treatments drugs like analapril or captopril. (I) has
CC excellent ability for producing optically active carboxylic acids from
CC carboxylic acid esters. Production of optically active carboxylic acids
CC by (I) is highly more selective, simple and environmentally favourable
CC when compared to conventional techniques. (I) is thermally stable at high
CC temperatures above 70 degrees Celsius. The present sequence represents a
CC carboxyl esterase amino acid sequence given in comparison with the
CC esterase of the present invention
XX
XX Sequence 312 AA;

Query Match 38.9%; Score 586.5; DB 4; Length 312;
Best Local Similarity 45.1%; Pred. No. 2.3e-53;
Matches 129; Conservative 54; Mismatches 100; Indels 3; Gaps 3;

QY 6 ANGITLEYEQGHRHHPHMLLMGLGQLDWPPEFIRGLAERGFVFCFDRDAGLSTK 65

DB 24 SNGLELHVEVGGNPDHPTILLIIMGLGQMLFWDFCKSLIDQGFYVIRFDNEDIGLSSK 83

QY 66 LEGV-KKPNRIARVFLASMG-L-KRPVPTLDDMALDTVGLMDALGIESTHVVGVSGMMI 123

DB 84 IRHKGRKRLNTKLMSRFTLGLNQGAPYTYLDMADVSLILEAMRIKKVNVIGASGMMI 143

QY 124 AQLGAKHGERVKSILMTITSSGNPRMPAPRQVQLKFMKRVKPSMDKEWIKYNLELLTT 183

DB 144 AQTIAAKYPEKVKALMTFTSSNQPLPPFPKQPLFSLIGKPKSDEDDGLINSLKLF 203

QY 184 IGSPGLDREKLAD-VKRSERCLCEGTORQLAAIIQSGSRVKLLRRIAVPTLVISGAE 242

DB 204 IGSPGVNVEATQARLLYQSRVHPAGVLOQFLAICTGSLQLDQKLSQPTLVINGSR 263

QY 243 DPLLPYOCGRDIADHPIGARFELIEGMDHPIPERHPIRLIELIAGH 288

DB 264 DRLPPSHGKAVAKAIGKARFELIQGMGHDIPPHFIPQLSYLFAH 309

RESULT 6
ABO73253
ID ABO73253 standard; protein; 422 AA.
XX

AC ABO73253;

XX 29-JUL-2004 (first entry)

XX *Pseudomonas aeruginosa* polypeptide #5428.

XX Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.

XX *Pseudomonas aeruginosa*.

OS US6551795-B1.

PN 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR N-PSDB; ABD06824.

XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 21999; 455pp; English.

XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a *P. aeruginosa* nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-*P. aeruginosa* drugs, as templates for recombinant
XX production of *P. aeruginosa*-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of *P. aeruginosa*-caused
XX infection, and in detection of *P. aeruginosa* sequences or other sequences
XX of *Pseudomonas* species using biotechnology. Sequences ABO67826-
XX ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

XX Sequence 422 AA;

Query Match 37.3%; Score 562.5; DB 7; Length 422;
Best Local Similarity 43.1%; Pred. No. 1.3e-50;
Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;

QY 4 VKANGITLEYEQGHRHHPHMLLMGLGQLDWPPEFIRGLAERGFVFCFDRDAGLS 63

DB 125 VELGDVRLAYQSIGRSDPALLIIMGLGQLHWPDEVVSALCEQGFVIRYDNRDVGLS 184

QY 64 TKLEGVKKPNRIARVFLASMG-L-KRPVPTLDDMALDTVGLMDALGIESTHVVGVSGMMI 123

DB 185 ANNVFVPSRLTYEVVRYRLGLFVSPAPYTLTDMAGDALHLLDALDIPQARVLGASGMMI 244

QY 124 AQLGAKHGERVKSILMTITSSGNPRMPAPRQVQLKFMKRVKPSMDKEWIKYNLELLTT 183

DB 245 AQHIADMAPQRLSLTTLVTSSGAGLPAPSSLLRLRLAR-REASREAOAVEQADLLAA 303

QY 184 IGSPGL--DREKALDVRKSIERCLCEGTORQLAAIIQSGSRVKLLRRIAVPTLVISGA 241

DB 304 LGSPEVRDDRQQLLQAARSYDRAFNPEGVQRLAILAEPKRVPLNRLNQPTLVINGT 363

QY 242 EDPPLPYOCGRDIADHPIGARFELIEGMDHPIPERHPIRLIELIAGHAAA 291

Db 364 ADPLLPMHGVHVAHNRGSELKLPGLAHRFQAEKPEKPIIAAVVPYLKA 413

RESULT 7

ADL05923
ID ADL05923 standard; protein; 360 AA.

AC
XX ADL05923;

DT 06-MAY-2004 (first entry)

DE M. catarrhalis protein #1689.

KW Moraxella catarrhalis; infection.

OS Moraxella catarrhalis.

PN US6673910-B1.

XX 06-JAN-2004.

PF 04-APR-2000; 2000US-00540236.

XX 08-APR-1999; 99US-0128416P.

PA (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL;

DR WPI; 2004-178127/17.

DR N-PSDB; ADL04003.

XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for preparing a composition for diagnosing, preventing or treating infection caused by Moraxella catarrhalis.

PS Disclosure; SEQ ID NO 3609; 429pp; English.

XX The invention relates to an isolated nucleic acid encoding an Moraxella catarrhalis polypeptide. The nucleic acid is useful for preparing a composition for diagnosing, preventing or treating infection caused by Moraxella catarrhalis. The present sequence represents the amino acid sequence of a M. catarrhalis protein.

XX Sequence 360 AA;

Query Match 32.08; Score 483; DB 8; Length 360;

Best Local Similarity 37.34; Pred. No. 3.1e-42; Mismatches 113; Indels 6; Gaps 3;

Matches 107; Conservative 61; Mismatches 113; Indels 6; Gaps 3;

QY 8 GITLEVEEQHRRHPSMLLMGLGGQLIDWPPEFIRGLAERGFVFCFNDRLAGLSTK-- 65

DB 42 GIKLVEAGNPEHPMIFITGLSQMFMSDFLKEPIDAGFVFIFDNRDTGLSKIQ 101

QY 66 LEGVKKPNIAVFLASMGKLPKPR---VPYTLDDMALDTVGLMDALGTESTHVGVSGMG 122

DB 102 IDGLPRLNTEFKMLKMQAGLSNRSEPVAYTLTDMAEDARLITKMLHNVNLIGASMGGM 161

QY 123 IAQILGNKGERVKSGLTMTTSSGNPNPAPRQVLOKFNVPKSMDSKKEWIKYNLELLT 182

DB 162 IAQIVARYPKYIKNLVLFSTNSRAFLRPNPKQFMTFVRPESHESRDMVRHSWNFMT 221

QY 183 TIGSPG-LDREKALDVKRSIERCLCEGTQORLAAILQSGSRVKLLRRIAAPTIVTISGA 241

DB 222 AVGSPGHLDTKGTFAIAEKYQNFHPLXVSOQTALASRSILRFTKQIRANTLVHGN 281

QY 242 EDPLLPYCCGRDIADHIPGARFELIEGMDHDIPERHIPRIELIAGH 288

DB 282 KDGIVAPNHCKILAKVISNARFVLVDGMGHDLPNYYPPYINGLISEH 328

RESULT 8

AAW30522

AAW30522 standard; protein; 267 AA.

AC
XX AAW30522;

DT 17-OCT-2003 (revised)

DT 26-OCT-1998 (first entry)

DE Kurthia sp. bioH gene-encoded polypeptide.

XX Biotin; vitamin; bioH gene; pimelyl CoA.

OS Kurthia sp. 538-KA26; (DSM 10609).

XX EP853127-A2.

XX 15-JUL-1998.

PF 18-SEP-1997; 97EP-00116237.

XX 27-SEP-1996; 96EP-00115540.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Furuichi Y, Hoshino T, Kimura H, Kiyasu T, Nagahashi Y;

DR WPI; 1998-364652/32.

DR N-PSDB; AAV42058.

XX New DNA and vectors encoding polypeptides - used for recombinant production of biotin.

XX Claim 1; Page 24-25; 45pp; English.

XX This is a polypeptide encoded by a newly isolated bioH gene (see AAV42058) of Kurthia sp. 538-KA26 (DSM 10609). This polypeptide is thought to be involved in the synthesis of the biotin precursor pimelyl CoA. Polypeptides (see AAW30518-25) encoded by novel bioD, bioA, bioP, bioH, bioH, bioH, bioH, bioH and bioC genes (see AAV42054-61) of Kurthia sp. are provided, as well as vectors comprising one or more of the genes, transformed cells, and a process for the production of biotin that comprises cultivating the transformed cells and isolating biotin from the culture medium. Biotin is used in the preparation of pharmaceutical, food or feed compositions. The synthesis is an improvement on prior art methods involving fermentation of e.g. E.coli and Bacillus sp. which have low productivity due to accumulation of DTB, a biotin precursor. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 267 AA;

Query Match 20.8%; Score 314.5; DB 2; Length 267;

Best Local Similarity 33.1%; Pred. No. 1.8e-24;

Matches 92; Conservative 43; Mismatches 104; Indels 39; Gaps 9;

QY 1 MAQVXKANGITLVEEQHRRHPSMLLMGLGGQLIDWPPEFIRGLAERGFVFCFNDRLA 60

DB 1 MPFVNHDSLEYEV--HGQGDPLLLINGLVNLSW-HRTVFTLAKR-FKLVFDNRGV 56

QY 61 GLSTKLEGVKKPNIAVFLASMGKLPKPRVPYTLDDMALDTVGLMDALGTESTHVGVSGM 120

DB 57 GKSS-----KPEQPSYIEMMAEDARAVLDVAVSDSAHVYGISMG 95

QY 121 GMIAQTILGAKHGERVKSGLTMTTSSGNPNPAPRQV---LQKFMVPKSMDSKKEW--- 173

DB 96 GMIAQRLAITYPVRSVSLVLCGTAGTTHIQSPSISTLMVSRASLTGSPRDNALAAP 155

QY 174 IKYNLELLTIGSPGLDREKALDVKRSIERCLCEGTQORLAAILQSGSRVKLLRIA 233

DB 156 IVYSQAFIEK--HPELIQI-----DIQRIEITPPSAVLSQLOACLTHDTSNE-LDKINI 208

QY 234 PTLVISGAEDPLLPYCCGRDIADHIPGARFELIEGMDH 271

DB 209 PTILIHGDADNLVPIYENGKMLAERIOGSGFHTVSCAGH 246

RESULT 9
ABM95684
ID ABM95684 standard; protein; 247 AA.
XX AC ABM95684;
XX DT 02-JUN-2005 (first entry)
XX DE M. xanthus protein sequence, seq id 14883.
XX KW Transgenic plant; DNA replication; gene regulation; gene expression.
XX OS Myxococcus xanthus.
XX PN US6833447-B1.
XX PD 21-DEC-2004.
XX PF 10-JUL-2001; 2001US-00902540.
XX PR 10-JUL-2000; 2000US-0217883P.
XX PA (MONS) MONSANTO TECHNOLOGY LLC.
XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX DR WPI; 2005-028716/03.
XX PT New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX PS Example 2; SEQ ID NO 14883; 25pp; English.
XX CC The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX SQ Sequence 247 AA;
Query Match 17.4%; Score 262.5; DB 9; Length 247;
Best Local Similarity 30.3%; Pred. No. 5.7e-19;
Matches 79; Conservative 38; Mismatches 107; Indels 37; Gaps 7;
QY 45 LAERGFRVCFDNRDAGLSTKLEGVKKPNIARVFLLASMGKPRVPYTLDDMALDVTGLM 104
DB 4 LARAGRPIRVYDHTGTS-----SHGVPGATPTVTLDDLAADVSVL 45
QY 105 DALGIETHVGVSMGMIAQILGAKHGERVKSITLM---ITSGNPRMPAPRPOVLOKF 161
DB 46 DGYGIERAHLVMSGLGLLQCLVALKYPERVLSITLISAQIFSPDPDFGMDPAVLAHF 105
QY 162 MRVP--KSMDEKWKYNLEL-LTTIGSP--GLDREKALDVRKSIERCLCPEGTOROLA 216
DB 106 ORAATLNSDEAEAGFQVLSRLCVGRARRSFDEARVARAVQDYRRALAPQ-----C 159
QY 217 AILQSGSRVKL-----LRRIVPTLVISGAEDPLLPYQCGRDIADHIPGARFELIEGMGH 271
DB 160 ALNHAGLSGGLGYGRTREIEAPLLVHGSVDPIVDAHGVALSRAVKGACLVTLHDAGH 219
QY 272 DIPERHPIRLIELIAGHAAA 292
DB 220 DLHPDDWETWTRAITAHTSAA 240

RESULT 10
ADI23937
ID ADI23937 standard; protein; 271 AA.
XX AC ADI23937;
XX DT 22-APR-2004 (first entry)
XX DE Streptomyces refuineus 024A locus ORF9 protein.
XX KW antimicrobial; fungicide; virucide; gene therapy; lipopeptide synthesis;
KW A54145; NRRL 3143; antimicrobial; antifungal; antiviral;
KW biosynthetic locus; 024A; ORF9.
XX OS Streptomyces refuineus.
XX PN US2003198981-A1.
XX PD 23-OCT-2003.
XX PF 24-DEC-2002; 2002US-00329079.
XX PR 26-DEC-2001; 2001US-0342133P.
XX PR 17-APR-2002; 2002US-0372789P.
XX PR 03-SEP-2002; 2002US-00232370.
XX PA (ECOP-) ECOPIA BIOSCIENCES INC.
XX PI Farnet CM, Staffa A, Zazopoulos E;
XX DR WPI; 2003-852784/79.
XX DR N-PSDB; ADI23938.
XX PT New isolated, purified or enriched nucleic acid, useful for synthesizing
PT lipopeptides, particularly from the biosynthetic locus A54145 and NRRL
PT 3143, that exhibits antimicrobial, antifungal or antiviral activity.
XX PS Example 4; SEQ ID NO 51; 69pp; English.
XX CC The invention describes an isolated, purified or enriched nucleic acid
CC (i) comprising: any of 4 fully defined sequences (SEQ. ID NO: 1, 6, 17
CC and 34), and their coding regions; a nucleic acid having at least 75%
CC sequence identity to a nucleic acid of (a); or complements of (a) or (b).
CC The methods and compositions of the present invention are useful for
CC synthesizing lipopeptides, particularly A54145 and NRRL 3143, exhibiting
CC antimicrobial, antifungal or antiviral activity. This is the amino acid
CC sequence of a protein encoded by the Streptomyces refuineus 024A (or NRRL
CC 3143) biosynthetic locus ORF9.
XX SQ Sequence 271 AA;
Query Match 15.1%; Score 227.5; DB 7; Length 271;
Best Local Similarity 26.9%; Pred. No. 3.5e-15;
Matches 84; Conservative 47; Mismatches 98; Indels 83; Gaps 13;
QY 1 MAQVKANGITLVEEEOGHRHPSMLLIGLGGOLIDWPEEFIRGLAERGFVFCFDRDA 60
DB 1.MPTTRINGALDHRGTG--SGPVLILINGSAKSAWHLHOVPALVAEGEATFTNR-- 56
QY 61 GLSTKLEGVKKPNIARVFLLASMGKLPKPR--VPYTLDDMALDVTGLMDALGIESHTVVGV 117
DB 57 -----GVPPSGGGPGFTLGDMAADTVGLIHLGIGFCAVVG 93
QY 118 SMGGMIAQILGAKHGERVKSITLMTSSGNPRMPA-----PRPQVLOK 160
DB 94 SLGARVAREVARTFDLVSRCLVAPRARSDRRAACTAAETALADSGVTLPPIRYAVR 153
QY 161 FWR--VPKSM--DKB--EWIKYNLELTTTIGSPGLDREKALDVRKSIERCLCPEGTORQ 214
DB 154 AMONLSPRTLADDRQIADWLDV-LELAADG-PGL-----RTOLE 191
QY 215 LAAILQSGSRVKLLRRIVPTLVISGAEDPLLPYQCGRDIADHIPGARFELIEGMGH--- 271

CC benzodioxole to sesamol. The Stenotrophomonas maltophilia bacteria
 CC contains a gene cluster (9320 nucleotide sequence given in the
 CC specification) that is involved in the conversion of 1,3-benzodioxole to
 CC sesamol. The method of the invention is useful for preparing sesamol
 CC which can be used as a synthetic raw material for pharmaceuticals such as
 CC antihypertensives. The present amino acid sequence represents a protein
 CC encoded by the Stenotrophomonas maltophilia sesamol production-related
 CC gene cluster of the invention.
 XX
 SQ Sequence 259 AA;
 Query Match 13.9%; Score 209.5; DB 8; Length 259;
 Best Local Similarity 28.9%; Pred. No. 2.7e-13;
 Matches 82; Conservative 36; Mismatches 109; Indels 57; Gaps 10;
 QY 4 VKANGITLEYEQGHRHPSMLLMGLGQIDW-----PEEFIRGLAERFVICED 56
 DB 13 VVANGIRTNTHDAG--DGAFLVMIHSGFGVSANWRLTMPE-----LATR-FRVIAPD 64
 QY 57 NRDAGLSTKLEGVKKNPNIARVFLASMGKLPVPYTYLDDMALDTVGLMDALGISTHVVVG 116
 DB 65 MVGFGYSQREPIH-----YSLDTWVQVALLDLDIEQASVVG 104
 QY 117 VSMGMIQILGAKHGERVKSILMITSSGNPRMPAPRPOVLQKFRVPSMDKEWIKY 176
 DB 105 NSFSGAIALALAIHPKRVKRLVLM-----GSGVSPFITTEGLDAVWGYQPSVE----- 153
 QY 177 NLELLTTIGSPGLDREKALDVRKSIERCLCEGTORQLAAIL-----QSG-----SRVKL 227
 DB 154 NRRALLDIFA--YDRKLNDLELAQMYKASIQGFQESFSAMPFPQNGVEMWASPIEQ 211
 QY 228 LRRIAVPTLVISGAEDPLPYQCGRIADHIPGARFELIEGMGH 271
 DB 212 IRGIEHQTLVHGREDKVIPLQNSYELLOAIQVPAQLHVFQKCGH 255

RESULT 15
 ID AAO22154
 XX AAO22154 standard; protein; 271 AA.
 AC AAO22154;
 XX
 DT 03-OCT-2002 (first entry)
 DE Ramoplanin biosynthetic ORF 9 protein.
 XX
 KW Ramoplanin; ramoplanin biosynthetic pathway polypeptide; antibiotic;
 KW biosynthesis gene cluster; bioengineering; peptide synthetase module;
 KW adenylation domain; hydroxyphenylglycine; HPG; antibiotic precursor;
 KW chlorinate; lipopeptide.
 XX
 OS Actinoplanes sp.
 XX
 PN WO200231155-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 15-OCT-2001; 2001WO-CA001462.
 XX
 PR 13-OCT-2000; 2000US-0239924P.
 PR 12-APR-2001; 2001US-0283296P.
 PR 24-JUL-2001; 2001US-00910813.
 XX
 PA (ECOP-) ECOPIA BIOSCIENCES INC.
 XX
 PI Farnet CW, Zazopoulos E, Staffa A;
 XX
 DR WPI: 2002-435445/46.
 DR N-PSDB; AAL40781.
 XX
 PT Novel isolated ramoplanin biosynthetic pathway polypeptide useful for
 PT chemically modifying biological molecule that is a substrate for a
 PT polypeptide encoded by a ramoplanin biosynthesis gene cluster.

XX
 PS Claim 14; Page 146-147; 212pp; English.
 XX
 CC The invention relates to an isolated ramoplanin biosynthetic pathway
 CC polypeptide selected from a polypeptide of open reading frames (ORF) 1-
 CC 32. The isolated polypeptides are useful for chemically modifying a
 CC biological molecule that is a substrate for a polypeptide encoded by a
 CC ramoplanin biosynthesis gene cluster, by contacting the biological
 CC molecule with the isolated polypeptide, where the polypeptide chemically
 CC modifies the biological molecule. The method comprises contacting the
 CC biological molecule with at least two different polypeptides encoded by
 CC ramoplanin ORFs 1-31. The polypeptides are useful for directing the
 CC biosynthesis of the antibiotic ramoplanin in microorganisms. An isolated
 CC gene cluster comprising the ORFs is useful as a substrate for
 CC bioengineering of antibiotic structures. An isolated polypeptide or its
 CC encoding nucleic acid sequence is useful for generating derivatives of
 CC ramoplanin, for improving production or for producing variants of other
 CC antibiotics of the peptide class. The isolated polypeptides are useful
 CC for synthesis of ramoplanin in vivo or in vitro, as an adenylation domain
 CC in conjunction with other peptide synthetase modules and allowing the
 CC incorporation of Thr into a peptide antibiotic precursor, for modifying
 CC fatty acid structure and/or enhancing fatty acid incorporation into the
 CC peptide antibiotic structure, for production of an hydroxyphenylglycine
 CC (HPG)-containing peptide antibiotic, for enhancing secretion of
 CC ramoplanin or its variants and derivatives, for enhancing uptake of
 CC precursors for ramoplanin biosynthesis, for enhancing production of
 CC ramoplanin products or its variants or derivatives, to chlorinate HPG of
 CC a peptide antibiotic precursor, and for designing specific nucleotide
 CC probes and primers for identifying and isolating putative lipopeptide
 CC producing microorganisms. This sequence represents one of the ORF
 CC proteins of the ramoplanin producing Actinoplanes sp. microorganism of
 CC the invention
 XX
 SQ Sequence 271 AA;
 Query Match 13.5%; Score 204; DB 5; Length 271;
 Best Local Similarity 25.0%; Pred. No. 1.1e-12;
 Matches 78; Conservative 38; Mismatches 120; Indels 76; Gaps 9;
 QY 4 VKANGITLEYEQGHRHPSMLLMGLGQIDWPEEFIRGLAERFVICFDNRDAGLS 63
 DB 9 VTTNGVRLAYRSGAGE--PVLIMNGSGSAGQTTVHTQTPALHEAGYSTVWFDSR----- 61
 QY 64 TKLEGVKKNPNIARVFLASMGKLPVPYTYLDDMALDTVGLMDALGISTHVVVSGMGI 123
 DB 62 -----GIPPSDV-----PAGKYSLABMTADTRGLIEALDILAPCRIVGTSLGAMI 105
 QY 124 AQILGAKHGERVKSILMITSSGNPRMPAPRQ-----VLQKFRVPSMDK----- 170
 DB 106 AQELAVDPHELVRCAVLATLARPDAARAAQADIDLLSGVTLPAAYEAATAVFMFS 165
 QY 171 -----EEWIKYNLELLTTIGSPGLDREKALDVRKSIERCLCEGTORQLAAILQ 220
 DB 166 PATLNDVAVREWLDI-FELSGTGVSA-----GQMAEL- 199
 QY 221 SGRVKLLRRRIAVPTLVISGAEDPLPYQCGRIADHIPGARFELIEGMGH-----DIPER 276
 DB 200 TGDRAALRSVTAPCKRVISFADDLITPPLHAAEVAEIPDCDLVEISRCGHLGYPDA 259
 QY 277 HIPRLIELIAGH 288
 DB 260 VNAAILFLDSH 271
 Search completed: April 13, 2006, 12:55:20
 Job time : 192 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 13, 2006, 12:55:38 ; Search time 41 Seconds
(without alignments)
689.945 Million cell updates/sec

Title: US-10-686-490D-2
Perfect score: 1509
Sequence: 1 MAQVKANGITLEYEQGHRH.....ERHPLRIELIAGHAAAE 294
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	589.5	39.1	321	1 S57530	carboxyl esterase
2	562.5	37.3	328	2 D83197	probable hydrolase
3	485	32.1	301	2 C70514	probable lipG prot
4	475.5	31.5	304	2 E87146	probable hydrolase
5	327.5	21.7	298	2 A57139	rdmc protein - Str
6	312.5	20.7	282	2 D95352	probable hydrolase
7	268.5	17.8	272	2 T36181	probable hydrolase
8	268.5	17.8	459	2 I60717	streptothricin-ace
9	234.5	15.5	265	2 A87193	probable hydrolase
10	229.5	15.2	288	2 C96028	probable 3-oxoadip
11	228.5	15.1	262	2 E70548	probable bpoC prot
12	227.5	15.1	276	2 T30594	conserved hypothet
13	217.5	14.4	261	2 T35708	hydrolase - Strept
14	211.5	14.0	256	2 AG3343	chloride peroxidase
15	211	14.0	282	2 S34609	carboxylesterase (
16	210	13.9	300	2 AG3214	cultivar specific
17	200.5	13.3	393	2 B87548	hypothetical prote
18	196.5	13.0	238	2 A69463	2-hydroxy-6-oxo-6-
19	196	13.0	283	2 B86504	poly(3-hydroxyalka
20	195	12.9	288	2 AB2770	hydrolase (impor
21	195	12.9	261	2 C97550	hydrolase (AJ27164
22	192	12.7	265	2 F83586	probable hydrolase
23	187	12.4	265	2 AB3115	beta-ketoacidate e
24	187	12.4	265	2 C98172	beta-ketoacidate e
25	186.5	12.4	275	2 F83243	probable hydrolase
26	184.5	12.2	647	2 C96834	unknown protein F5
27	182	12.1	263	2 C83616	beta-ketoacidate e
28	182	12.1	614	2 E95568	unknown protein, 1
29	180	11.9	285	2 S29306	poly(3-hydroxyalka

30 178 11.8 259 2 E84125 hydrolase BH3805 (

31 176.5 11.7 272 2 C87711 hydrolase, alpha/b

32 176.5 11.7 375 2 T35015 probable 3-oxoadip

33 176 11.7 283 2 B87546 acetoin dehydrogen

34 175 11.6 286 2 G70948 hypothetical prote

35 174 11.5 370 2 B83128 probable hydrolase

36 172.5 11.4 300 2 A70616 probable epnF prot

37 171 11.3 269 2 E87595 hydrolase, alpha/b

38 168 11.1 330 2 G87401 epoxide hydrolase

39 167.5 11.1 278 2 H75259 epoxide hydrolase

40 167 11.1 262 2 A89831 hypothetical prote

41 165.5 11.0 267 2 AE3589 3-oxoacidate enol-

42 165.5 11.0 286 1 C35124 2,6-dioxo-6-phenyl

43 165.5 11.0 296 2 F95380 probable hydrolase

44 165 10.9 270 2 AE3187 hypothetical prote

45 165 10.9 322 2 B70957 probable epha prot

ALIGNMENTS

RESULT 1

S57530
carboxyl esterase - Acinetobacter calcoaceticus
C:Species: Acinetobacter calcoaceticus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S57530
R:Kok, R.G.; Bart, A.; Hellingwerf, K.J.
submitted to the EMBL Data Library, June 1995
A:Description: Characterization of the estBR operon of Acinetobacter calcoaceticus BD4.
A:Reference number: S57529
A:Accession: S57530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <KOK>
A:Cross-references: UNIPROT:Q57172; UNIPARC:UPI00001727AA; EMBL:X88895
C:Genetics:
A:Start codon: GTG
C:Superfamily: carboxyl esterase

Query Match 39.1%; Score 589.5; DB 1; Length 321;
Best Local Similarity 45.5%; Pred. No. 1.5e-41;
Matches 130; Conservative 53; Mismatches 100; Indels 3; Gaps 3;

QY 6 ANGITLEEEOGRRHPSMLLGLGGQLIDPEEFIRGLAERGFVFCFNDNRLGSLTK 65
DB 33 SNGLELHVEVGNDPDHPTILLMGLGAOMLFWPDPFCKSLIDQGFVIRFDRDIGSSK 92
QY 66 LSGV-KKPNIAVFLASML-KPRVPYTLDDMDALTVGLMDALGIESTHVGVSGMI 123
DB 93 IRHKGKRLNTLKMRSFTLGLNGQCAPYTLVMAEDVSLLEAMRIKKVNVIGASMGMI 152
QY 124 AOLGAKHGERVKSJTLMTSSGNRMPAPRQVLQKFMKVPKSMOKBEWIKYNLELTT 183
DB 153 AQLIAAKYPERVEKALMTSTNNQPLPPFPKQLFSLGKPKSDEDDGIINSLKLFEL 212
QY 184 IGSPGLDREKLALD-VRKSIERCLCEPQORQLAAILOSGRVKLLRRIAVPTLVISGAE 242
DB 213 IGSPGVNQVEAIQTARKLYQRSYHPAGVLQQLFALICTGSLLOKLOKISQPTLVINGSR 272
QY 243 DPLLPPQCRDADHIDPGARFELIEMGHIDIPERHIFRILIELIAGH 288
DB 273 DRLLPPSHGKAVAKAKGAKFELIQMGHDIPPHFIPQLSYLFAHH 318

RESULT 2

D83197
probable hydrolase PA3586 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 12-Jul-2004
C:Accession: D83197
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Li
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Li

.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:2043737; PMID:10984043
A;Accession: D83197
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-328 <STO>
A;Cross-references: UNIPROT:Q9HY40; UNIPARC:UPI000000CSA79; GB:AE004779; GB:AE004091; NID:10984043
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3586
C;Superfamily: tropinesterase

Query Match 37.3%; Score 562.5; DB 2; Length 328;
Best Local Similarity 43.1%; Pred. No. 2.7e-39;
Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;

QY 4 VKANGITLEYEQGRRHHPMSMLINGLGGQLIDWPPEFIRGLAERGFVFCFNDAGLS 63
DB 31 VELGDVRLAQSGRSDPALLVWGLGGQLIHPDVEVSALCEQGFVIRYDNRDGLS 90
QY 64 TKLEGVKKPNRIARVFLASGLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGMGI 123
DB 91 AMNVVPSSRLTYEVVRYRLGLPVSAFYTLTDMAGDALHLLDALDIPQAHVGLASGMGI 150
QY 124 AOLGAKHGERVKSLLTMITSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYNLELLTT 183
DB 151 AOHADMAPORLLSLTVMTSSGAEGLPAPSESLRLRLAR-REAAAREQAVEQADLLAA 209
QY 184 IGSFGL--DREKALDVRKSIERCLCEGTQORQLAAIQLQSGSRVKKLRRVAVPTLVISGA 241
DB 210 LGSFVDRDDQQQLLQARSYDRAFNEGVQRQLLALAEPSRVPLNRLQVTLVHGT 269
QY 242 EDPLLPYQCGRDIAHPIGARFELIEGMDHPIPERHIPRIELIAGHAAA 291
DB 270 ADPLLPVHGWHVAAHIRGSELKIPGLAHRFOEAPKEPLAAVVPYLKA 319

RESULT 3
C70614
probable lipG protein - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70614
R;Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70614
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-301 <COL>
A;Cross-references: UNIPROT:P96935; UNIPARC:UPI0000165242; GB:Z92772; GB:AL123456; NID:98295987
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: lipG
C;Superfamily: carboxyl esterase

Query Match 32.1%; Score 485; DB 2; Length 301;
Best Local Similarity 40.1%; Pred. No. 6.9e-33;
Matches 112; Conservative 50; Mismatches 111; Indels 6; Gaps 4;

QY 9 ITLEYEQGRRHHPMSMLINGLGGQLIDWPPEFIRGLAERGFVFCFNDAGLSKLEG 68
DB 13 VKLYEDMGDLDPVPLVLLINGLGAQLWRTFCARLVAKGLRVIRYDNRDGLSKYTER 72
QY 69 VKKPN--IARVFLASGLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGMGI 126
DB 73 -HFGQPLATVRLVSMGLPSQAAYTLEDMAADAAALLDLDVKAHVVGASGMGI 131

QY 127 LGAKHGERVKSLLTMITSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYN-LELLTTIG 185
DB 132 FAARFAQRKTLLAVIFSSNNHRLPAPPAPALLALLTGPDPSPROVDVNAVRSKIIG 191
QY 186 SPG--LDREKALDVRKSIERCLCEGTQORQLAAIQLQSGSRVKKLRRVAVPTLVISGAED 243
DB 192 SPAYPIPEQVRAEAAESYDRNPHFWGIAQOQFSAIILGSGSLRYDRRIVAPTIVHGRAD 251
QY 244 PLLPYOCGRDIADHIPGARFELIEGMDHPIPERHIPRLI 282
DB 252 KLMRPFGRGAVARAINGARLVLDLIDGMGHDLPLQLMDRVI 290

RESULT 4
E87146
probable hydrolase [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87146
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ieam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Nature 409, 1007-1011, 2001
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: E87146
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-304 <STO>
A;Cross-references: UNIPROT:Q9CBK5; UNIPARC:UPI000000CE03; GB:AL450380; NID:gl3093568;
C;Genetics:
A;Gene: lipG
C;Superfamily: carboxyl esterase

Query Match 31.5%; Score 475.5; DB 2; Length 304;
Best Local Similarity 36.9%; Pred. No. 4.3e-32;
Matches 109; Conservative 54; Mismatches 117; Indels 15; Gaps 4;

QY 9 ITLEYEQGRRHHPMSMLINGLGGQLIDWPPEFIRGLAERGFVFCFNDAGLSKLEG 68
DB 13 LKLYEDMGNDVDPVPLVLLINGLGAQLWRTAFCEKLVAGQLRVYRDNDRDGLSRTDS 72
QY 69 VKKP-----NIARVFLASGLKPRVPTLDDMALDTVGLMDALGIESTHVGVSGM 120
DB 73 TEOCPSPQLTARLIRFWL-----GQKNACAYLTEDMTDDAVALDHLSTIERAHIVGASMG 128
QY 121 GMAIQTILGAKHGERVKSLLTMITSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYNLEL 180
DB 129 GMAIQIFARFFTRSLAVFFSSNNRPLPPAPALLALLTGPPTGSRDDVVVDNVVR 188
QY 181 LTTI-GSP--GLDRKALDVRKSIERCLCEGTQORQLAAIQLQSGSRVKKLRRVAVPTLV 237
DB 189 VTKITGSPLYRMPPEQVTRTAAEYDRSFYPLGVSRQFSAILGSGSLHYNQRIIAPTIV 248
QY 238 IESGADPLPYOCGRDIADHIPGARFELIEGMDHPIPERHIPRIELIAGHAAA 292
DB 249 IHGRADKLVPSGSGRAVARAITCARLVLPFDGMGHDLFPQQLMDQAVGLVMSNFAKA 303

RESULT 5
A57139
rnc protein - Streptomyces purpurascens
C;Species: Streptomyces purpurascens
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 09-Jul-2004
C;Accession: A57139
R;Niemi, J.; Maentzela, P.
J. Bacteriol. 177, 2942-2945, 1995
A;Title: Nucleotide sequences and expression of genes from *Streptomyces purpurascens* t
A;Reference number: A57139; MUID:95270621; PMID:7751313
A;Accession: A57139
A;Status: preliminary

Qy	61	GLSTKLEGVCKPNATARVPELLASMLCKRVP-YTLDDMALDVTGLMDALGISTHSTVVGVSM	119
Db	61	GLSTK-----YPPGQGYAFDNADVVRVLGVRISAAHVGVMSL	101
Qy	120	GGMAIQILGAKGHERVSKSLTLMITS---SGNPRMPAPRQVQLQKPMRVPKPSMDKEWIKY	176
Db	102	GGMIGQATALKHPERVLSLTATSSPGVMNTHLPASGTAWMDHNMVEVDSWDAEAVAY	161
Qy	177	NLE---LLTTIGSPGLDREKALDVRKSIERCLCPG-----TORQLAAILQSGSRVKLL	228
Db	162	MLEDARLAVSTVHPFDDEAE-----TRAFIERDFRSGYLSATNHSVLFEISDAWQDR-L	215
Qy	229	RRIAVPLTVTISGAEDDLPLPQCGRDIDAHHPGARFELIEGMGHDIPERHIPRLIELI	285
Db	216	PEMKVPLLVLTIGTADPVFVPEHGAATAVDGARLVEIIEGGHELHPADWDKISAI	272

```

RESULT 7
T36181
probable hydrolase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36181
K:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajan
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21600
A:Accession: T36181
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-272 <SAU>
A:Cross-references: UNIPROT:Q9Z4Z8; UNIPARC:UPI000000DAF2A; EMBL:AL035707; PTDN
A:Experimental source: strain A3(2)

```

```
C;Superfamily: carboxyl esterase
```

	Query Match Best Local Similarity Matches	17.8%; Score 268.5; DB 2; Pred.No. 6.7e-15; Conservative 44;	Mismatches 109; Indels 79; Gaps 10
Qy	1 MAOVKANGITLEYEQ----	CHRRHPSMILLIMGLGGQLIDWPESFIRCLASRGFRVCFD	56
	: :	: :: :	:
Dd	1 MPVLTVNGIRINYTDAPPAGANAPAVLLVMGSGSGRAHLHQVPALVAAGFRVISFD		60
Qy	57 NRDAGLSTLKEGVKKPNIARVFLASNLGRVPTYTLDDMALDTVGMDALGIESTHYVG		116
	: :	: :: :	:
Dd	61 NRIAPSEPCP-----	-GFGIDDLVADTAALVEERLRUGPCR VAG	99
Qy	117 VSMGMIAOTILGAHGSRVKSLTIMITSSGNP-----	--RMPAPRPQV	157

```

RESULT 8
I60717
streptothricin-acetyl-transferase - Escherichia coli
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C:Accession: I60717
R:Rietze, E.; Brevet, J.
Plasmid 25, 217-220, 1991

```

Db	98	VARPELVAALMATRGRLDRTRQFFHAAAEAFHDSGIGQLPSGYNAKVRLLNLSR--KT 155
Qy	168	MDKE-----EWI-KYNLELLTTIGSPCLDRKLDLVKRSIERCLCPGEGTQRLAAILQSG 222

probable bpoC protein - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: E70548
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.B.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70548
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <COL>
A;Cross-references: UNIPROT:O06420; UNIPARC:UPI000000D5E6A; GB:Z95558; GB:AL123456; NID:9
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: bpoC
C;Superfamily: tropinesterase

Query Match 15.1%; Score 228.5; DB 2; Length 262;
Best Local Similarity 28.1%; Pred. No. 1.4e-11;
Matches 81; Conservative 29; Mismatches 103; Indels 75; Gaps 10;

QY 9 ITLEYEQGRRHPSMLLIGLQGLIDWPPEFIRGLAERGFRVICDNRDAGLSTKLEG 68
DB 2 INLAYDNG--TGPVVFIAAGRGAGRTWHPQVPAFLAAGYRCITFDNRGIGATENAEG 59
QY 69 VKKPNIARVFLASMGKLPKRPVYTLDDMALDTVGLMDALGIESTHVVGSMGMIAQILG 128
DB 60 -----FTQTWADTAALIE TLDTAPARVVGSMGAFIAQELM 97
QY 129 AKGGERVKSLLMIT-----SSGNRMPA---PPQVLOKFMKVPKS 167
DB 98 VVAPELVSSAVLMATRGRLDRARQFNKAERLYDSGVQLPPTVDARARLENFSR--KT 155
QY 168 MDKE---EMIKYNLELLTTIGSGPLDREKALDVRKSIERCLCEGTORQLAAILQSGS 223
DB 156 LNDVAVGDVIAW-FSWWPIKSTPL-----RC-----QLDCAPQT-N 191
QY 224 RVKLLRIAVPTLVISGAEDPLPYQCGRIADHIPGAREFELIEGMGH 271
DB 192 RLPAYRNAAPVLVIGPADDVTPPYLGRVADALPNRYLQIPDAGH 239

RESULT 12
T30594
conserved hypothetical protein PCZA361.30 - Amycolatopsis orientalis
C;Species: Amycolatopsis orientalis
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T30594
R;Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard, N. Chem. Biol. 3, 155-162, 1998
A;Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin g
A;Reference number: Z18804
A;Accession: T30594
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-276 <VAN>
A;Cross-references: UNIPROT:O52809; UNIPARC:UPI000000BA4D0; EMBL:AJ223998
C;Superfamily: carboxyl esterase

Query Match 15.1%; Score 227.5; DB 2; Length 276;
Best Local Similarity 27.3%; Pred. No. 1.8e-11;
Matches 84; Conservative 39; Mismatches 130; Indels 55; Gaps 9;

QY 1 MAQKANGITLVEEQGRRHPSMLLIGLQGLIDWPPEFIRGLAERGFRVICFNDA 60
DB 1 MLMTTETGRLSYHD--HGDGSPVLLTGTGAPSSVDLHQVPLVRAAGFRVITMDNR-- 56
QY 61 GLSTKLEGVKKPNIARVFLASMGKLPKRPVYTLDDMALDTVGLMDALGIESTHVVGSMG 120
DB 57 GIPPSDEGT-----GFTIDDLVADVAALIEHLGVAPCRVGTSMG 97
QY 121 GMIAQILGAKHGERVKSLLMITSSGNRMPAPRPQVLOKFMKVPKSMDEIKYNLEL 180

C;Accession: E70548
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.B.; Taylor, K.; Whitehead, S.; Barrrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70548
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <COL>
A;Cross-references: UNIPROT:O06420; UNIPARC:UPI000000D5E6A; GB:Z95558; GB:AL123456; NID:9
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: bpoC
C;Superfamily: tropinesterase

Query Match 15.1%; Score 217.5; DB 2; Length 261;
Best Local Similarity 27.4%; Pred. No. 1.1e-10;
Matches 69; Conservative 27; Mismatches 67; Indels 89; Gaps 7;

QY 49 GFRVICFDRDAGLSTKLEGVKKPNIARVFLASMGKLPKRPVYTLDDMALDTVGLMDALG 108
DB 50 GIRAISVDQDSCIT-----VNPVPVTPPEVLADLDLDDALG 88
QY 109 IESTHVVGSMGMIAQILGAKHGERVKSLLMITSSGNRMPAPRPQVLOKFMKVPKSM 168
DB 89 LARAHLLGTSFGGVAQAARHPRKVASLVLVATTTPSYAMGSAIDELLE-----M 140
QY 169 DKEEMIKYNLELLTTIGSGPLDREKALDVRKSIERCLCEGTORQLAAILQSGSRVK-- 226
DB 141 SHE-----DROQANADY-----FFTPEG-----QAGQARPA 167
QY 227 --LLRR-----IAVPTLVISGAEDPLPYQCGRIADHIPGA 261
DB 168 GYLTRTPQSVRRHDAARRHEIRDGLGGITAPTLVHGTDRLAPVEGALLMERRIPNA 227
QY 262 RPELIEGMGHDI 273
DB 228 ELCPIEGGRHGI 239

RESULT 14
AG3343
chloride peroxidase (EC 1.11.1.10) [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 12-Jul-2004
C;Accession: AG3343
R;DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete.
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melite
A;Reference number: AD3252; PMID:11756688
A;Accession: AG3343
A;Status: preliminary

Db 98 SYIAQELALAHPELLDVAVLMACG-----RSSLVQRVL-----AEGAKLIELCTEL 145
QY 181 -----LTTIGSGPLDREKLA---LDVRKSIERCLCEGTORQLAAILQSGSRVKL 227
DB 146 PPGYLAARAHMNLGPATLADDDLTGDLWDLFASDNWNGPVRAQLQLSAL---PDRIDA 202
QY 228 LRIAIVPTLVISGAEDPLPYQCGRIADHIPGAREFELIEGMGH---DIPERHIPBLIE 283
DB 203 YRAIKVPCHVIFSEFHDLVAPPSAGRELAAAIATHRTIPGCGHFGYLENPEAVNRELLR 262
QY 284 LIAGHAAA 291
DB 263 FIRAESAA 270


```

DR Pfam; PF00561; Abhydrolase_1; 1.
SQ SEQUENCE      312 AA; 34681 MW; 14E020BA7654D905 CRC64;

Query Match          39.1%; Score 589.5; DB 2; Length 312;
'Best Local Similarity 45.5%; Pred No. 8.6e-37;
Matches 130; Conservative

Qy    6 ANGITLEYEQGHRHPHSMLLINGLGGQLDWPEEFIRGLAERGFRVICFDNRDAGLSTK 65
     :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   24 SNGLEHLHVEVGGNPDPHTILLINGLGQAQLFWPDFCKSLIDQGFIYREDNRDIGLSKK 83
     ::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   66 LEGV-KKPNTARVFLLASWCL-KPRVPYTLDMDALDTVLGMALGIESTHVGVSGMGMI 123
     :|::||::||::||::||::||::||::||::||::||::||::||::||
Db   84 IRHKGRKLNTLKMSRFTLLGNGQGAIPPTLYDMAEDVSLLLEAMRIKVNVIGASMGGMI 143
     ||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   124 AQTLAGKHGBRVKSLTLMITSSGNPMPPAPPOVOLQFMRPVKSMDEEWIKYNLELLTT 183
     ||::||::||::||::||::||::||::||::||::||::||::||::||
Db   144 AQIIAAKYPKVEKIALMFTSNNOPLLPPFPFKOLFSLICKPKSSEDDGIINHSKLFEI 203
     ||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   184 IGSFGPLDREKIALD-VRKSIERLCCEPGEQRQAAILQSGRVKLRRAIVPTLVISGAE 242
     ||::||::||::||::||::||::||::||::||::||::||::||::||
Db   204 IGSPGYVNVEAIQTARKLYQRSYHPAGVLOQQFLLAILCTGSLQLDKIQSOPTLVHGSR 263
     ||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   243 DPLLPQCGRDIADHPICGARFELIEGHGDIPERHIPRLIELTAGH 288
     ||::||::||::||::||::||::||::||::||::||::||::||::||
Db   264 DRLLPSHKGAVAKAYKAIGAKFELIOGNGHDIPPFFIPQLSYFAHH 309
     ||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 4
Q57172 ACTAD PRELIMINARY; PRT; 312 AA.
ID Q57172 ACTAD PRELIMINARY; PRT; 312 AA.
AC Q57172;
AT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE EstB protein (Putative hydrolase).
DS Name=estB; OrderedLocusNames=ACIADI064;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX MEDLINE=95400495; PubMed=7670642;
RA Gaisdoerfer W., Frosch C.S., Haspel G., Ehrt S., Hillen W.;
RT "Two genes encoding proteins with similarities to rubredoxin and
RT rubredoxin reductase are required for conversion of dodecane to lauric
RT acid in Acinetobacter calcoaceticus ADP1.";
RL Microbiology 141:1425-1432(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX MEDLINE=97228433; PubMed=9074511; DOI=10.1016/S0378-1119(96)00728-7;
RA Gaisdoerfer W., Ratajczak A., Hillen W.;
RT "Nucleotide sequence of a putative periplasmic Mn superoxide dismutase
RT from Acinetobacter calcoaceticus ADP1.";
RL Gene 186:305-308(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX MEDLINE=98162050; PubMed=9501429;
RA Gaisdoerfer W., Ratajczak A., Hillen W.;
RT "Transcription of ppk from Acinetobacter sp. strain ADP1, encoding a
RT putative polyphosphate kinase, is induced by phosphate starvation.";
RL Appl. Environ. Microbiol. 64:896-901(1998).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RA Gaisdoerfer W., Kok R.G., Ratajczak A., Hellingwerf K.J., Hillen W.;
RT "The genes rubA and rubB for alkane degradation in Acinetobacter sp.
RT strain ADP1 are in an operon with estB, encoding an esterase, and
```

DR Pfam: PF00561; Abhydrolase_1; 1.
KW Hydrolase.
SQ SEQUENCE 312 AA; 34681 MW; 14E020BA7654D905 CRC64;
Query Match 39.1%; Score 589.5; DB 2; Length 312;
Best Local Similarity 45.5%; Pred. No. 8.6e-37;
Matches 130; Conservative 53; Mismatches 100; Indels 3; Gaps 3;
QY 6 ANGITLEYEQGHRHPSMLLMGLGQGLIDWPPEFIRGLAERGFVFCFDRDAGLSTK 65
DB 24 SNGLELHVEVGNGPDHPTILLMGLGQMLFWDPFCKSLDQGFVYRFDNRDGLSSK 83
QY 66 LEGV-KKPNIAKRVFLASMG-L-KRPVPTLDDMALDVTGLMDALGISTHVVGVSMGMI 123
DB 84 IRHKGRKLTNKLMSRFTLGLNGQCAPYTLDYMAEDVSLLENRIKKVNVIGASMGMI 143
QY 124 AOILGAKHGERVKSLLTMITSSGNPRMPAPRPOVLOKFMRPVKMSDKKEWIKYNLELLTT 183
DB 144 AOIIAAKPEKVEKALMFTSNNQPLPPPPKQLFSLIGKPKSDEDEGINHSLKLF 203
QY 184 IGSPGLDREKLALD-VRKSIERCCLPECTQRLAAILQSGSRVKLLRRIAVPTLIVISGAE 242
DB 204 IGSPGVNVQVEAIQTARKLYQKSYHPAGVLOQFLAILCTGSLQLDQKISOPTLVINGSR 263
QY 243 DPLLPGQGRDIADHIPGARELIEGMDHDPERHIFRLIELIAGH 288
DB 264 DRLLPPSHGKAVAKAIGAKFELIQGMGHDIPPHFIPQLSYLFAHH 309
RESULT 4
Q57172 ACIAD PRELIMINARY; PRT; 312 AA.
AC Q57172;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE BstB protein (Putative hydrolase).
GN Name=estB; OrderedLocusNames=ACIAD1064;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX MEDLINE=95400495; PubMed=7670642;
RA Geissdorfer W., Frosch C.S., Haspel G., Ehart S., Hillen W.;
RT "Two genes encoding proteins with similarities to rubredoxin and
rubredoxin reductase are required for conversion of dodecane to lauric
acid in Acinetobacter calcoaceticus ADP1.";
RL Microbiology 141:1425-1432 (1995).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX MEDLINE=97228433; PubMed=9074511; DOI=10.1016/S0378-1119(96)00728-7;
RA Geissdorfer W., Ratajczak A., Hillen W.;
RT "Nucleotide sequence of a putative periplasmic Mn superoxide dismutase
from Acinetobacter calcoaceticus ADP1.";
RL Gene 186:305-308 (1997).
RN (3)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX MEDLINE=98162050; PubMed=9501429;
RA Geissdorfer W., Ratajczak A., Hillen W.;
RT "transcription of pbk from Acinetobacter sp. strain ADP1, encoding a
putative polyphosphate kinase, is induced by phosphate starvation.";
RL Appl. Environ. Microbiol. 64:896-901 (1998).
RN (4)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RA Geissdorfer W., Kok R.G., Ratajczak A., Hellingwerf K.J., Hillen W.;
RT "The genes rubA and rubB for alkane degradation in Acinetobacter sp.
strain ADP1 are in an operon with estB, encoding an esterase, and
estC, encoding an esterase, and
estD, encoding an esterase, and
estE, encoding an esterase, and
estF, encoding an esterase, and
estG, encoding an esterase, and
estH, encoding an esterase, and
estI, encoding an esterase, and
estJ, encoding an esterase, and
estK, encoding an esterase, and
estL, encoding an esterase, and
estM, encoding an esterase, and
estN, encoding an esterase, and
estO, encoding an esterase, and
estP, encoding an esterase, and
estQ, encoding an esterase, and
estR, encoding an esterase, and
estS, encoding an esterase, and
estT, encoding an esterase, and
estU, encoding an esterase, and
estV, encoding an esterase, and
estW, encoding an esterase, and
estX, encoding an esterase, and
estY, encoding an esterase, and
estZ, encoding an esterase, and
estAA, encoding an esterase, and
estAB, encoding an esterase, and
estAC, encoding an esterase, and
estAD, encoding an esterase, and
estAE, encoding an esterase, and
estAF, encoding an esterase, and
estAG, encoding an esterase, and
estAH, encoding an esterase, and
estAI, encoding an esterase, and
estAJ, encoding an esterase, and
estAK, encoding an esterase, and
estAL, encoding an esterase, and
estAM, encoding an esterase, and
estAN, encoding an esterase, and
estAO, encoding an esterase, and
estAP, encoding an esterase, and
estAQ, encoding an esterase, and
estAR, encoding an esterase, and
estAS, encoding an esterase, and
estAT, encoding an esterase, and
estAU, encoding an esterase, and
estAV, encoding an esterase, and
estAW, encoding an esterase, and
estAX, encoding an esterase, and
estAY, encoding an esterase, and
estAZ, encoding an esterase, and
estBA, encoding an esterase, and
estBB, encoding an esterase, and
estBC, encoding an esterase, and
estBD, encoding an esterase, and
estBE, encoding an esterase, and
estBF, encoding an esterase, and
estBG, encoding an esterase, and
estBH, encoding an esterase, and
estBI, encoding an esterase, and
estBJ, encoding an esterase, and
estBK, encoding an esterase, and
estBL, encoding an esterase, and
estBM, encoding an esterase, and
estBN, encoding an esterase, and
estBO, encoding an esterase, and
estBP, encoding an esterase, and
estBQ, encoding an esterase, and
estBR, encoding an esterase, and
estBS, encoding an esterase, and
estBT, encoding an esterase, and
estBU, encoding an esterase, and
estBV, encoding an esterase, and
estBW, encoding an esterase, and
estBX, encoding an esterase, and
estBY, encoding an esterase, and
estBZ, encoding an esterase, and
estCA, encoding an esterase, and
estCB, encoding an esterase, and
estCC, encoding an esterase, and
estCD, encoding an esterase, and
estCE, encoding an esterase, and
estCF, encoding an esterase, and
estCG, encoding an esterase, and
estCH, encoding an esterase, and
estCI, encoding an esterase, and
estCJ, encoding an esterase, and
estCK, encoding an esterase, and
estCL, encoding an esterase, and
estCM, encoding an esterase, and
estCN, encoding an esterase, and
estCO, encoding an esterase, and
estCP, encoding an esterase, and
estCQ, encoding an esterase, and
estCR, encoding an esterase, and
estCS, encoding an esterase, and
estCT, encoding an esterase, and
estCU, encoding an esterase, and
estCV, encoding an esterase, and
estCW, encoding an esterase, and
estCX, encoding an esterase, and
estCY, encoding an esterase, and
estCZ, encoding an esterase, and
estDA, encoding an esterase, and
estDB, encoding an esterase, and
estDC, encoding an esterase, and
estDD, encoding an esterase, and
estDE, encoding an esterase, and
estDF, encoding an esterase, and
estDG, encoding an esterase, and
estDH, encoding an esterase, and
estDI, encoding an esterase, and
estDJ, encoding an esterase, and
estDK, encoding an esterase, and
estDL, encoding an esterase, and
estDM, encoding an esterase, and
estDN, encoding an esterase, and
estDO, encoding an esterase, and
estDP, encoding an esterase, and
estDQ, encoding an esterase, and
estDR, encoding an esterase, and
estDS, encoding an esterase, and
estDT, encoding an esterase, and
estDU, encoding an esterase, and
estDV, encoding an esterase, and
estDW, encoding an esterase, and
estDX, encoding an esterase, and
estDY, encoding an esterase, and
estDZ, encoding an esterase, and
estEA, encoding an esterase, and
estEB, encoding an esterase, and
estEC, encoding an esterase, and
estED, encoding an esterase, and
estEE, encoding an esterase, and
estEF, encoding an esterase, and
estEG, encoding an esterase, and
estEH, encoding an esterase, and
estEI, encoding an esterase, and
estEJ, encoding an esterase, and
estEK, encoding an esterase, and
estEL, encoding an esterase, and
estEM, encoding an esterase, and
estEN, encoding an esterase, and
estEO, encoding an esterase, and
estEP, encoding an esterase, and
estEQ, encoding an esterase, and
estER, encoding an esterase, and
estES, encoding an esterase, and
estET, encoding an esterase, and
estEU, encoding an esterase, and
estEV, encoding an esterase, and
estEW, encoding an esterase, and
estEX, encoding an esterase, and
estEY, encoding an esterase, and
estEZ, encoding an esterase, and
estFA, encoding an esterase, and
estFB, encoding an esterase, and
estFC, encoding an esterase, and
estFD, encoding an esterase, and
estFE, encoding an esterase, and
estFF, encoding an esterase, and
estFG, encoding an esterase, and
estFH, encoding an esterase, and
estFI, encoding an esterase, and
estFJ, encoding an esterase, and
estFK, encoding an esterase, and
estFL, encoding an esterase, and
estFM, encoding an esterase, and
estFN, encoding an esterase, and
estFO, encoding an esterase, and
estFP, encoding an esterase, and
estFQ, encoding an esterase, and
estFR, encoding an esterase, and
estFS, encoding an esterase, and
estFT, encoding an esterase, and
estFU, encoding an esterase, and
estFV, encoding an esterase, and
estFW, encoding an esterase, and
estFX, encoding an esterase, and
estFY, encoding an esterase, and
estFZ, encoding an esterase, and
estGA, encoding an esterase, and
estGB, encoding an esterase, and
estGC, encoding an esterase, and
estGD, encoding an esterase, and
estGE, encoding an esterase, and
estGF, encoding an esterase, and
estGG, encoding an esterase, and
estGH, encoding an esterase, and
estGI, encoding an esterase, and
estGJ, encoding an esterase, and
estGK, encoding an esterase, and
estGL, encoding an esterase, and
estGM, encoding an esterase, and
estGN, encoding an esterase, and
estGO, encoding an esterase, and
estGP, encoding an esterase, and
estGQ, encoding an esterase, and
estGR, encoding an esterase, and
estGS, encoding an esterase, and
estGT, encoding an esterase, and
estGU, encoding an esterase, and
estGV, encoding an esterase, and
estGW, encoding an esterase, and
estGX, encoding an esterase, and
estGY, encoding an esterase, and
estGZ, encoding an esterase, and
estHA, encoding an esterase, and
estHB, encoding an esterase, and
estHC, encoding an esterase, and
estHD, encoding an esterase, and
estHE, encoding an esterase, and
estHF, encoding an esterase, and
estHG, encoding an esterase, and
estHH, encoding an esterase, and
estHI, encoding an esterase, and
estHJ, encoding an esterase, and
estHK, encoding an esterase, and
estHL, encoding an esterase, and
estHM, encoding an esterase, and
estHN, encoding an esterase, and
estHO, encoding an esterase, and
estHP, encoding an esterase, and
estHQ, encoding an esterase, and
estHR, encoding an esterase, and
estHS, encoding an esterase, and
estHT, encoding an esterase, and
estHU, encoding an esterase, and
estHV, encoding an esterase, and
estHW, encoding an esterase, and
estHX, encoding an esterase, and
estHY, encoding an esterase, and
estHZ, encoding an esterase, and
estIA, encoding an esterase, and
estIB, encoding an esterase, and
estIC, encoding an esterase, and
estID, encoding an esterase, and
estIE, encoding an esterase, and
estIF, encoding an esterase, and
estIG, encoding an esterase, and
estIH, encoding an esterase, and
estII, encoding an esterase, and
estIJ, encoding an esterase, and
estIK, encoding an esterase, and
estIL, encoding an esterase, and
estIM, encoding an esterase, and
estIN, encoding an esterase, and
est

```

oxyR.";
J. Bacteriol. 181:2925-2928 (1999).
[5]
NUCLEOTIDE SEQUENCE.
PubMed=15514110; DOI=10.1093/nar/gkh910;
Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT ADP1, a versatile and naturally transformation competent bacterium."
Nucleic Acids Res. 32:5766-5779 (2004).
EMBL; Z46863; CA86927.1; -; Genomic DNA.
EMBL; CR543861; CAG67952.1; -; Genomic DNA.
PR; S57330; S57330.
GO; GO:0016787; F:hydrolase activity; IEA.
InterPro; IPR000073; A/b hydrolase.
PFam; PF00561; Abhydrolase_1; 1.
Complete proteome; Hydrolase.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 312 AA; 34661 MW; 14E020BA7654D905 CRC64;

Query Match 39.1%; Score 589.5; DB 2; Length 312;
Best Local Similarity 45.5%; Pred. No. 8.6e-37;
Matches 130; Conservative 53; Mismatches 100; Indels 3; Gaps 3;

QY 6 ANGTLLEYBOGHHHPHSMMLIMGLGQIDWPEEFIRGLAERGVICFDRDAGLSTK 65
DB 24 SNGLEHVEVGNPDHPTILLIMGLGQMLFWDFFCKSLIDQGYVIREFDRDGLSSK 83

QY 66 LEGV-KKPNIAARVFLASMG-L-KRPVPTLDDMALDTVGLMDALGIESTHVGVSMGMI 123
DB 84 IRHKGRILNTKLMSRTTGLNGQAPYTYLDMAEDVSLLEAWRIKKNVIGASMGMI 143

QY 124 AQLGAKHGERVKSITLMTSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYNLELLTT 183
DB 144 AQLIAAKYKPKVKALMTSSNQLPPPPKQLPSLICKPKSDDEGIINSLKLF 203

QY 184 IGSPGLDREKALD-VKRSIERCLCEGTORQLAAILOSRSVKLLRIAPVTLVSGAE 242
DB 204 IGSPGVYQVEAIQTARKLYQSRYPAGVLQQLFALICTGSLQLQDKQISOPTLVINGSR 263

QY 243 DPLPYQCGRDIAHDPGARFELTEGMGHDIPIERHPIRLIELIAGH 288
DB 264 DRLPPSHGKAVAKAIGKARFELIQGMGHDIPIPHFIPQLSYLFAHH 309

RESULT 5
Q6D508_ERWCT PRELIMINARY; PRT; 306 AA.
AC Q6D508;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Probable hydrolase.
GN OrderedLocusNames=SCA2233;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Ball K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Fraser A., Hance Z., Hauser H., Jagers K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
carotovora subsp. atroseptica and characterization of virulence
factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110 (2004).

oxyR.";
J. Bacteriol. 181:2925-2928 (1999).
[5]
NUCLEOTIDE SEQUENCE.
PubMed=15514110; DOI=10.1093/nar/gkh910;
Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT ADP1, a versatile and naturally transformation competent bacterium."
Nucleic Acids Res. 32:5766-5779 (2004).
EMBL; Z46863; CA86927.1; -; Genomic DNA.
EMBL; CR543861; CAG67952.1; -; Genomic DNA.
PR; S57330; S57330.
GO; GO:0016787; F:hydrolase activity; IEA.
InterPro; IPR000073; A/b hydrolase.
PFam; PF00561; Abhydrolase_1; 1.
Complete proteome; Hydrolase.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 312 AA; 34661 MW; 14E020BA7654D905 CRC64;

Query Match 38.4%; Score 579; DB 2; Length 306;
Best Local Similarity 43.9%; Pred. No. 5.3e-36;
Matches 122; Conservative 45; Mismatches 105; Indels 6; Gaps 3;

QY 4 VKANGITLVEEQGHRHHPHSMMLIMGLGQIDWPEEFIRGLAERGVICFDRDAGLS 63
DB 5 IRKGVNIAYESFDESTEIIILVAGLSOMISWSAFCTVVSRRGVRVRFDRDVGCS 64

QY 64 TKLEGVKKPNIAARVFLASMG-L-KRPVPTLDDMALDTVGLMDALGIESTHVGVSMGMI 123
DB 65 SHFYDSPVFNIGEVNKAISANRQPDVPTLDDVAGDLGLMDALSITAAHFVGRSMGMI 124

QY 124 AQLGAKHGERVKSITLMTSSGNPRMPAPRQVLOKFMVPSKMDKEWIKYNLEL 180
DB 125 AQLIAARHPERVLSTLIIMSSGNPSLPOTAPDVM-ALMTQAPAPNPFLEPEYLNHRL 183

QY 181 LTTIGSP--GLDREKALDVRKSIERCLCEGTORQLAAILOSRSVKLLRIAPVTLVI 238
DB 184 AERLACKIYFFDTERYSQFKEELRECFNPEGTROMAALVASGDRKKLLATIAAPTLVI 243

QY 239 SGAEDPLLPYQCGRDIAHDPGARFELTEGMGHDIPIER 276
DB 244 HGTDPLFPVSHGEDVAINIPEAEFNLIDGMGHDIPIQ 281

RESULT 6
Q9HY40_PSEAE PRELIMINARY; PRT; 328 AA.
AC Q9HY40;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable hydrolase.
GN OrderedLocusNames=PA3586;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reiter J., Sater M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen."
RL Nature 406:959-964 (2000).
EMBL; AE004779; AAG06974.1; -; Genomic DNA.
PR; D83197; D83197.
GO; GO:0016787; F:hydrolase activity; IEA.
GO; GO:0016740; F:transferase activity; IEA.
InterPro; IPR000073; A/b hydrolase.
PFam; PF00561; Abhydrolase_1; 1.
Complete proteome; Hydrolase.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 328 AA; 35604 MW; C3DDA69157505F77 CRC64;

Query Match 37.3%; Score 562.5; DB 2; Length 328;
Best Local Similarity 43.1%; Pred. No. 1.1e-34;
Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;

```

```

QY 4 VKANGITLEYEQGRRHHPMSLLIMGLGQLIDWPPEFIRGLAERGFRVCFDNRDAGLS 63
Db 31 VELGDVRLAYQSIGRDSFALLVWGLGGQLIHWPDDEVVSCQGRFVRIRYDNRDVGLS 90
QY 64 TKLEGVKKPNIAKVFLASMGKPRVPYTLDDMALDVTGLMDALGIBSTHVGVSMGMI 123
Db 91 ANWVPVSSRLTYEVVRYLGLPVSAPTYTLDMAGDALHLLDALDIPQAHVIGASMGMI 150
QY 124 AQILGAKHGERVKSITLMTSSGNPRMPAPQVQLQKPMRVKSMDEWIKYNLELTT 183
Db 151 AQHIAADNAPORLLSLLTMTSSGAEGLPAPESLRLRLAR-REASREQAEEQQADLLA 209
QY 184 IGSPLG--DREKALDVRKSIERCPCGEGTORQALAAIILQSGSRVKLLRRIAVPTLVISGA 241
Db 210 LGSFEVRDRQLLQAAARSVDRAFNPGEVQRLALIAEFSRVPLNRLQVPTLVHGT 269
QY 242 EDPLLPVQCGRDIADHIFGARFELIEGMGHDIPIERHIFPLTIELIAGHAAA 291
Db 270 ADPLLPVHGVHVAHIRGSELKLIPLGLAHRFQAEKPELLAAVVPYLKA 319

RESULT 7
Q88N9_PSEPK
ID Q88N9_PSEPK PRELIMINARY; PRT; 378 AA.
AC Q88N9;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE Hydrolyase, alpha/beta fold family.
GN OrderedLocusNames=PP1064;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfle E.K., Scanlan D., Tran K.,
RA Meazzaz A., Utterback T.R., Rizzo M., Lee K., Kosack D., Mestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eiben J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016778; AAN66689.1; -; Genomic_DNA.
DR TIGR; PP1064; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000379; Ser_gstrs.
DR Pfam; PF00561; Abhydrolase_1; I.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 378 AA; 41119 MW; 055BAB0F98F3EBBA CRC64;

Query Match 37.3%; Score 562.5; DB 2; Length 378;
Best Local Similarity 44.4%; Pred. No. 1.3e-34;
Matches 122; Conservative 47; Mismatches 103; Indels 3; Gaps 2;

QY 3 QVANGITLEYEQGRRHHPMSLLIMGLGQLIDWPPEFIRGLAERGFRVCFDNRDAGL 62
Db 78 RAELGAVSLVYQSVGAPRDPALLVWGLGGQLIHWPDDEVVSCQGRFVRIRYDNRDVG 137
QY 63 STKLEGVKKPNIAKVFLASMGKPRVPYTLDDMALDVTGLMDALGIBSTHVGVSMGMI 122
Db 138 SRWQVPPHANLTLELRYKLGIPVSAPTYTLDMAGDALGILMDALGVQRPHVIGSMGM 197
QY 123 IAQILGAKHGERVKSITLMTSSGNPRMPAPQVQLQKPMRVKSMDEWIKYNLELTT 182

```

```

Db 198 IAQHLAAMAPERVRSITLVSSSGAGLPAPDPALVOLLAR-RSAPPREVAIEQQADLLA 256
QY 183 TIGSPGL--DREKALDVRKSIERCPCGEGTORQALAAIILQSGSRVKLLRRIAVPTLVISG 240
Db 257 ALGSEPVDRDDREVLLHQAAQAYDRAFNPQGAQKQIMAILAEPSPRVLLNQLRVPTLVVHG 316
QY 241 AEDPLLPVQCGRDIADHIFGARFELIEGMGHDIPIE 275
Db 317 TADPLLPVHGVHVAHIRGSELKLIPLGLAHRFQAE 351

RESULT 8
Q4K713_PSEFS
ID Q4K713_PSEFS PRELIMINARY; PRT; 397 AA.
AC Q4K713;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Hydrolyase, alpha/beta fold family.
GN ORFNames=PFL4889;
OS Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PF-5;
RX PubMed=15980861; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodi D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Winn M.B., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouri H.M., Pierson E., Pierson L. III, Thomasow L., Loper J.;
RT "Complete genome sequence of the plant commensal Pseudomonas
RT fluorescens Pf-5.";
RL Nat. Biotechnol. 23:873-878(2005).
DR EMBL; CP000076; AAY94119.1; -; Genomic_DNA.
KW Hydrolase.
SQ SEQUENCE 397 AA; 42869 MW; 2F77FPD2116F6E86 CRC64;

Query Match 37.2%; Score 561.5; DB 2; Length 397;
Best Local Similarity 45.4%; Pred. No. 1.6e-34;
Matches 122; Conservative 43; Mismatches 101; Indels 3; Gaps 2;

QY 9 ITLEYEQGRRHHPMSLLIMGLGQLIDWPPEFIRGLAERGFRVCFDNRDAGLSKLEG 68
Db 94 VSLAYQSIGRDSFALLVWGLGGQLIHWPDDEVVSCQGRFVRIRYDNRDVG 153
QY 69 VKKPNIAKVFLASMGKPRVPYTLDDMALDVTGLMDALGIBSTHVGVSMGMIQILG 128
Db 154 PVSANLTTFEVLRYKLGIPVSAPTYTLDMAGDALGILMDALHIEQFHVIGSMGMIAQHLA 213
QY 129 AKHGERVKSITLMTSSGNPRMPAPQVQLQKPMRVKSMDEWIKYNLELTTIGSPG 188
Db 214 AMEPRVESLTLMTSSGAEGLPAPNAALVQLLSR-RNAPSREAAEQADLLAALGSPK 272
QY 189 L--DREKALDVRKSIERCPCGEGTORQALAAIILQSGSRVKLLRRIAVPTLVISGADPLL 246
Db 273 VSDRQVLLQAAVAYDRAFNPGEVQRLALIAEFSRVALLKQLRVPTLVHGTADPLL 332
QY 247 PYQCGRDIADHIFGARFELIEGMGHDIPIE 275
Db 333 PVHGVHVAHIRGSELKLIPLGLAHRFQAE 361

RESULT 9
Q97TK4_CLOAB
ID Q97TK4_CLOAB PRELIMINARY; PRT; 299 AA.
AC Q97TK4;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

```

```
DE Carboxyl esterase, a/b hydrolase.
GN OrderedLocusNames=CAP0097;
QS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RV NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 824 / DSM 792 / VRM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R.L., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
RL EMBL; AB001438; AAK76842.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR003089; AB_hydrolase.
DR InterPro; IPR000379; Ser_estra.
DR Pfam; PF00561; Abhydrolase_1; 1.
DR PRINTS; PR00111; ABHYDROLASE.
KW Complete proteome; Hydrolase; Plasmid.
SQ SEQUENCE 299 AA; 33303 MW; DD958146552C420E CRC64;

Query Match 36.9%; Score 556.5; DB 2; Length 299;
Best Local Similarity 41.1%; Pred. No. 2.7e-34;
Matches 118; Conservative 53; Mismatches 113; Indels 3; Gaps 2;

Qy 1 MAQVKGITLEYEEQGRHHPHSMGLGQQLIDWPEFIRGLAERGRVFCFNRDA 60
Db 1 MAHVNTGQIQEYVFGKRTNPTVLVLAGNQAQNFWSDFCEMLAQNQVIRFNRDA 60

Qy 61 GLSTKLGKGVKPNRIARVFLASMLGKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMG 120
Db 61 GLSTKFDPAAGIPDSMKIYQAQEQKPKTATYLEDMAVDVAGLLDALEIKKAKICGASG 120

Qy 121 GMTAIIQAKHGERVKSILMITSSGNRMPAPRPQVLQKPMRVKSMDEKWIKNLEL 180
Db 121 GMTAIIQAKHGERVKSILMITSSGNRMPAPRPQVLQKPMRVKSMDEKWIKNLEL 179

Qy 181 LTTIGSPG--LDREKALDVKSTIERCLCEGTORQALAILQSGSRVKLLRRVAVPTLVI 238
Db 180 WKLLWSKGFPEERAIYTESYDRCPYQGAVRQNAALVANGDRKRLSLLRVPTLVI 239

Qy 239 SGAEDPLLPVOCGRDIADHIPGARFELIEGMDHIDPERHIPRIELIAGHA--AAEA 294
Db 240 HGTADPLIPVEAGKDTARTIPNAKLLIEGMDHIDPERHIPRIELIAGHA--AAEA 286

RESULT 10
Q89R88 BRAJA PRELIMINARY; PRT; 304 AA.
ID Q89R88 BRAJA PRELIMINARY; PRT; 304 AA.
AC Q89R88 BRAJA PRELIMINARY; PRT; 304 AA.
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B12884 protein.
GN OrderedLocusNames=b12884;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiacae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RV NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneo T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,

RA Saemoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000840; BAC48149.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser_estra.
DR Pfam; PF00561; Abhydrolase_1; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 33095 MW; 58552CC1F028B6A8 CRC64;

Query Match 36.6%; Score 552.5; DB 2; Length 304;
Best Local Similarity 44.6%; Pred. No. 5.6e-34;
Matches 133; Conservative 42; Mismatches 112; Indels 11; Gaps 4;

Qy 4 VKANGITLEYEEQGRHHPHSMGLGQQLIDWPEFIRGLAERGRVFCFNRDAGLS 63
Db 10 VRANGIDICVEIFGNDNAEPFLLLIMGLGAQMIHWDADFCEQLAVHGFVRIRFNRDICKS 69

Qy 64 TKLEGVKKPNRIARVFLASMLGKPRVPYTLDDMALDTVGLMDALGIESTHVGVSMGMI 123
Db 70 SHLTGGRKLTPTLELLKRLRIPVAATYKLIIDWARTVGLMDALGIKSAHLVGASMGMI 129

Qy 124 AQILGAKHGERVKSILMITSSGNRMPAPRPQVLQKPMRVKSMDEKWIY--NLELL 181
Db 130 AQEVTLSFPFVRSLTSSIMSTTGNRVPPTTREAAMLMAPPKSKKEEFVRYQGTWKVL 189

Qy 182 TITGSGLDREKALD---VRKSIERCLCEGTORQALAILQSGSRVKLLRRVAVPTLVI 238
Db 190 RAGAFP---EEBALPDRAERVFARGLNPAGVGRQLRAVLASGSRKERLHAKMTPLVI 245

Qy 239 SGAEDPLLPVOCGRDIADHIPGARFELIEGMDHIDPERHIPRIELIAGHA--AAEA 294
Db 246 HGTVDPLVPEGGKDTAASIPGAKLLMVGMDHIDPERHIPRIELIAGHA--AAEA 303

RESULT 11
Q87XK0 PSISM PRELIMINARY; PRT; 330 AA.
ID Q87XK0 PSISM PRELIMINARY; PRT; 330 AA.
AC Q87XK0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hydrolase, alpha/beta fold family.
GN OrderedLocusNames=PSPT04178;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RV NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Daviden T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RL Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AB016853; AAO57634.1; -; Genomic DNA.
DR TIGR; PSPT04178; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser_estra.
```

[illegible]

```

QY 239 SGAEDELLPQCCRDIAHPIPGARFELIEGMGHDIPE 275
DB 267 HGTADPLLPVGHVAAHQSGQLRLIPGLAHRFQ 303

RESULT 14
Q4J684 AZOVI PRELIMINARY; PRT; 328 AA.
AC Q4J684;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DE Alpha/beta hydrolase fold precursor.
GN ORFNames=AvindRAFT 8586;
OS Azotobacter vinelandii AvOP.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=322710;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AvOP.
RG US DOE Joint Genome Institute (JGI-PGP);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Piliuck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Azotobacter vinelandii
RT AvOP.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AvOP;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Azotobacter vinelandii
RT AvOP.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AvOP;
RG US DOE Joint Genome Institute;
RA Hammon N., Israni S., Piliuck S., Richardson P.;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAU0300001; EAM06907.1; -; Genomic_DNA.
KW Hydrolase; Signal.
FT SIGNAL
SQ SEQUENCE 328 AA; 36068 MW; 3258374E36CC4A7D CRC64;

Query Match 35.6%; Score 536.5; DB 2; Length 328;
Best Local Similarity 43.9%; Pred. No. 1e-32;
Matches 118; Conservative 44; Mismatches 104; Indels 3; Gaps 2;

QY 9 ITLEYEQGHRHPSMLLMGLGGQLIDWPEEFIRGLAERFVICFDRDAGLSTKLEG 68
DB 36 VQLAYOSIGRETPALVNLVWGLGGQLIDWPEVLGLCQGFVIRFDRDVLGSAWRQ 95
QY 69 VKKPNIAVFLASMLGKRPVPTLDDMALDTVGLMDALGIESTHVGVSMGGMIAQILG 128
DB 96 VPKVSLTYGVLRLGLPVSAPYGLRDMARDITLALMARLQVPRFHVGLVSMGGMIAQHWA 155
QY 129 AKHGERVKSILMITSSGNPRMPAPRQVLOKFMVPKSKMEWKYNLELLTTIGSP- 187
DB 156 DLAPERVRSLLTMITSSALGLPGPSALL-KLLAQREAGSREVAIERQVELLAUSPQ 214
QY 188 -GLDREKIALDVRSKIERCLCPBGTQRLAAIILQSGSRVKLLRRIAVPTLVISGAEDPL 246

215 VVDREQLRQAAATAYDRAFNPEGVKRQILAVLAEPSEVLLERLHLPTLVHGTADPLL 274
QY 247 PYQCCGRDIAHPIPGARFELIEGMGHDIPE 275
DB 275 PVMHGVHVAHVHVGSELRLIPGLAHRFQ 303

RESULT 15
Q582C7 9TRYP PRELIMINARY; PRT; 312 AA.
AC Q582C7;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DE Alpha/beta hydrolase fold family, putative.
GN ORFNames=TB927.5.2380;
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Fai G., Peterson J.,
RA Khalak H.G., Salberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091655; AAX80442.1; -; Genomic_DNA.
KW Hydrolase.
SQ SEQUENCE 312 AA; 34433 MW; ECH85ECF9E41DE87 CRC64;

Query Match 32.3%; Score 488; DB 2; Length 312;
Best Local Similarity 39.0%; Pred. No. 4.9e-29;
Matches 110; Conservative 54; Mismatches 114; Indels 4; Gaps 3;

QY 9 ITLEYEQGHRHPSMLLMGLGGQLIDWPEEFIRGLAERFVICFDRDAGLSTKLEG 68
DB 21 ISLSYDTFGNSKXPCCLLVGLAGVGRVWRDAFCEMIAKKGFVVRVYDRDVLGSLTHLN 80
QY 69 VKKPNIAVFLASMLGKRPVPTLDDMALDTVGLMDALGIESTHVGVSMGGMIAQILG 128
DB 81 QTPNVNQCLLPQFLSFLRKVPTLDDMAADGNLLTALGIERAHVGVSMGGMIAQIMA 140
QY 129 AKHGERVKSILMITSSGNPRMPAPRQVLOKFMVPKSKMEWKYNLELLTTIGSPG 188
DB 141 IKYPSRVSRLGIYHSGTSSKRVPTFTSKLLFMKKPKSAALEDVVDVDFKALARHFRPG 200
QY 189 --LDREKIALDVRSKIERCL-CPEGTQRLAAIILQSGSRVKLLRRIAVPTLVISGAEDPL 245
DB 201 YNVDEEERFKLAQEQLRANDYPQGMRLQALAILSAKSECECLKTITPTLIHGMDEL 260
QY 246 LPYQCGRIDADIH-IGARFELIEGMGHDIPEIRHIELIA 286
DB 261 VPYQNGLOIAEAVGPAKLVITYPRMGHEIPVLMFMSISOEIA 302

Search completed: April 13, 2006, 12:59:18
Job time : 234 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 13, 2006, 12:59:33 ; Search time 47 Seconds
(without alignments)
517.163 Million cell updates/sec

Title: US-10-686-490D-2
Perfect score: 1509
Sequence: 1 MAQVANGTLEYEQHRH.....ERHPLIELIAGHAAABA 294

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCUS_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607.5	40.3	322	2	US-09-328-352-5739
2	562.5	37.3	422	2	US-09-252-991A-21999
3	483	32.0	360	2	US-09-540-236-3609
4	314.5	20.8	267	2	US-08-935-263-10
5	314.5	20.8	267	2	US-09-594-185-10
6	314.5	20.8	267	2	US-10-033-078-10
7	314.5	20.8	267	2	US-10-763-933-10
8	262.5	17.4	247	2	US-09-902-540-14883
9	192	12.7	287	2	US-09-252-991A-22466
10	189.5	12.6	251	2	US-09-345-469-4
11	189	12.5	268	2	US-09-252-991A-32242
12	185	12.3	265	2	US-09-902-540-11639
13	183	12.1	289	2	US-09-469-211A-6
14	183	12.1	289	2	US-09-446-681-6
15	182.5	12.1	271	2	US-09-902-540-12298
16	181.5	12.0	264	2	US-09-134-001C-5592
17	180	11.9	338	2	US-09-252-991A-32574
18	175	11.6	261	2	US-09-902-540-13754
19	174	11.5	334	2	US-09-902-540-11426
20	174	11.5	786	2	US-09-252-991A-32352
21	173.5	11.5	271	2	US-09-328-352-7546
22	168	11.1	281	2	US-09-902-540-11020
23	163	10.8	287	2	US-09-902-540-15478
24	162.5	10.8	275	2	US-09-902-540-15896
25	157	10.4	320	2	US-10-272-490-6
26	156.5	10.4	274	2	US-09-252-991A-21590
27	155.5	10.3	286	2	US-09-355-166-9

28	155.5	10.3	388	1	US-08-232-519-2	Sequence 2, Appli
29	155.5	10.3	388	1	US-08-456-956-2	Sequence 2, Appli
30	155	10.3	320	2	US-09-252-991A-18301	Sequence 18301, A
31	150	9.9	327	2	US-10-272-490-58	Sequence 58, Appl
32	149.5	9.9	274	2	US-09-949-016-6477	Sequence 6477, Ap
33	149.5	9.9	279	2	US-09-949-016-7659	Sequence 7659, Ap
34	149	9.9	326	2	US-10-272-490-20	Sequence 20, Appl
35	147	9.7	321	2	US-09-252-991A-29844	Sequence 29844, A
36	146	9.7	269	2	US-10-272-490-22	Sequence 22, Appl
37	146	9.7	462	2	US-09-712-363-209	Sequence 209, App
38	144.5	9.6	331	2	US-10-272-490-50	Sequence 50, Appl
39	144.5	9.6	339	2	US-10-272-490-60	Sequence 60, Appl
40	144	9.5	276	2	US-09-489-039A-13894	Sequence 13894, A
41	143	9.5	296	2	US-09-489-039A-7531	Sequence 7531, Ap
42	141.5	9.4	321	2	US-08-909-125-5	Sequence 5, Appli
43	138.5	9.2	312	2	US-09-424-349A-5	Sequence 2, Appli
44	138.5	9.2	554	1	US-08-106-761-2	Sequence 2, Appli
45	138.5	9.2	554	2	US-08-909-125-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-328-352-5739
; Sequence 5739, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5739
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5739

Query Match	40.3%	Score	607.5	DB 2	Length	322	
Best Local Similarity	46.9%	Pred. No.	3.1e-57	Indels	5	Gaps	4
Matches	134	Conservative	49	Mismatches	98		
QY	7	NGITLEYEEQHRHPSMLLIMGLGQITDWPFEPIRGLAERGFVFCFNDRDAGLSTKL	66				
DB	33	NGIELHVEVGKPEHPTILLINGLCAQMLFWDFFCKSLIDQGFVIRFDNRDGLSSKV	92				
QY	67	--EGVKPNIRAVFLLASMGKPR-VPYTLDMDALDTVGLMDALGIESTHVGVSMGMI	123				
DB	93	RHOG-KRLNTWKLMGRFALGRNQAPYTYLDMADDVSMLLDRLGVSKARVIGASMGMI	151				
QY	124	AQILGAKGHRVKSJTLMTSSGNRPMAPRQVQLQKMRVPKSMDEKWKYNLELLTT	183				
DB	152	AQILAATPEKVEKLGMLFTNNQPLPPKQLLSLGKPESDERDEGINVHSLKLFOL	211				
QY	184	IGSPG-LDREKALDVRSIERCLCEGTQRLAAAILQSGSRVKLRRIAVPTLVISGAE	242				
DB	212	IGSPGYINHIEAVQARKLYQRSYYPAGVQLQFLLCTGSLQLDREIKQPTLVLHGR	271				
QY	243	DPLLPYQGRDIADHIPGARFELIEGMGHDPERHPIRLIELIAGH	288				
DB	272	DRLLPPSHGKAVAKAISGAKFELIDMGHDPHPAIPQLSGLFAHH	317				

RESULT 2
US-09-252-991A-21999
; Sequence 21999, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252.991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 21999
;; LENGTH: 422
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21999

Query Match 37.3%; Score 562.5; DB 2; Length 422;
Best Local Similarity 43.1%; Pred. No. 3.5e-52;
Matches 125; Conservative 45; Mismatches 117; Indels 3; Gaps 2;

QY 4 VKANGITLVEEQQHRRHPSMLLMGLGGQLIDWPPEFIRGLAERGRVFCFNRDAGLS 63
DB 125 VELGDVRLAYQSGRSDPALLVMGLGGQLIHPDDEVVSALCEQGGFRVIRYDNRDVGLS 184

QY 64 TKLEGVKKPNIRARVFLASMLKPRVPTLDDMALDTVGLMDALGIESTHVGVSMGGM 123
DB 185 ANVPVPSRLTYEVVRYRLGLPVSAFYTLTDAGDALHLLDALDIPQAHVGLASMGGM 244

QY 124 AQILGAKHGRVSKLTMTTSSGNPRMPAPRQVLOKFMKVPKSMKDEWIKYNLELLT 183
DB 245 AQIADMAPQLLSLTVMTSSAGELPAPSESILLRLAR-REAAAREQAVEQQADLAA 303

QY 184 IGSFGL--DREKALDVRKSIERCLCEGTQOROLAAILOSGSRVKLLRRIAVPTLVISGA 241
DB 304 LGSFVDRDDQQLLQARSYDRANFEGVQORQLALAEPSVPLNRLQVPTLVHGT 363

QY 242 EDPLLPYQCGRDIADHIPGARFELIEGMGHDIPERHIPRLIELIAGHAAA 291
DB 364 ADPLLPVHGVAHNRGSELKLIPLGLAHRFQEAPEKPELLAAVVPYKA 413

RESULT 3
US-09-540-236-3609
; Sequence 3609, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3609
; LENGTH: 360
; TYPE: PRT
; ORGANISM: M.catarrhalis
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (250)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-540-236-3609

Query Match 32.0%; Score 483; DB 2; Length 360;
Best Local Similarity 37.3%; Pred. No. 1.1e-43;
Matches 107; Conservative 61; Mismatches 113; Indels 6; Gaps 3;

QY 8 GITLVEEQQHRRHPSMLLMGLGGQLIDWPPEFIRGLAERGRVFCFNRDAGLSTK-- 65
DB 42 GIKLCVAGNPEHPMPFITGLSQMFWSDQFLKRFIDAGFFVIRFNRDGLSKSQ 101

QY 66 LEGVKKPNIRARVFLASMLKPR---VPTLDDMALDTVGLMDALGIESTHVGVSMGGM 122
DB 102 IDGLPRLNTPKMLKMQAGLSNRPVAYTLTDMAEDAAARLKTMTQLHNVLIGASMGGM 161

QY 123 IAQILGAKHGRVSKLTMTTSSGNPRMPAPRQVLOKFMKVPKSMKDEWIKYNLELLT 182
DB 162 IAQIVAAIRPKYIKNUVLVLFSTNSNRAFLRPPNPQPMTFVRRPESHSDMVRHSVWFMT 221

QY 183 TIGSPG-LDREKALDVRKSIERCLCEGTQOROLAAILOSGSRVKLLRRIAVPTLVISGA 241
DB 222 AVGSPGHLDIKGTPTAAEKYQYRNFHPLXVSQQLTALASRSILRFTKQIRANTLVIHGN 281

QY 242 EDPLLPYQCGRDIADHIPGARFELIEGMGHDIPERHIPRLIELIAGH 288
DB 282 KDGLVAPNKGKILAKVISNARFVLVDGMGHDLPNYPYPLINGLISEH 328

RESULT 4

US-08-935-263-10
; Sequence 10, Application US/08935263A
; Patent No. 6117669
; GENERAL INFORMATION:
; APPLICANT: Furuichi, Yasuhiro
; APPLICANT: Hoshino, Tatsuo
; APPLICANT: Kimura, Hitoshi
; APPLICANT: Kiyasu, Tatsuya
; APPLICANT: Nagahashi, Yoshie
; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
; FILE REFERENCE: Biotin Genes
; CURRENT APPLICATION NUMBER: US/08/935,263A
; CURRENT FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: EP 96115540.5
; EARLIER FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Kurthia sp.
US-08-935-263-10

Query Match 20.8%; Score 314.5; DB 2; Length 267;
Best Local Similarity 33.1%; Pred. No. 1.3e-25;
Matches 92; Conservative 43; Mismatches 104; Indels 39; Gaps 9;

QY 1 MAQVKANGITLVEEQQHRRHPSMLLMGLGGQLIDWPPEFIRGLAERGRVFCFNRDA 60
DB 1 MPFVADNESLYEV--HCGGDPDLLLIMGLGYNLSW-HRTVPTLAKR-FKVIIVFNRGV 56

QY 61 GLSTKLEGVKKPNIRARVFLASMLKPRVPTLDDMALDTVGLMDALGIESTHVGVSMG 120
DB 57 GKSS-----KPEQFYSIEMMAEDARAVLDVAVSDSAHYVIGISM 95

QY 121 GMIAQILGAKHGRVSKLTMTTSSGNPRMPAPRQV---LQKFMKVPKSMKDEW--- 173
DB 96 GMIAQRLAITYPERVRSVLGCTTAGTTHIQSPPEISTLMVSRASLTGSPRDNALAAP 155

QY 174 IKYNLELLTIGSPGLDREKALDVRKSIERCLCEGTQOROLAAILOSGSRVKLLRRIAV 233
DB 156 IVYSQAFIEK--HPELIQE---DIQKRIEITTPPSAYLSQLQACLTHTDSNE-LDKINI 208

QY 234 PTLVTSAGBDDPLLPYQCGRDIADHIPGARFELIEGMGH 271
DB 209 PTLIIHGADNLVPYENGKMLAERIQQSQFHTVSCAGH 246

RESULT 5

US-09-594-185-10
; Sequence 10, Application US/09594185
; Patent No. 6365388
; GENERAL INFORMATION:
; APPLICANT: Furuichi, Yasuhiro
; APPLICANT: Hoshino, Tatsuo
; APPLICANT: Kimura, Hitoshi
; APPLICANT: Kiyasu, Tatsuya
; APPLICANT: Nagahashi, Yoshie

QY 1 MAQVKANGITLEYEQGRRHPSMLLMGLGGQLIDWPPEFIRGLAERGFVICFDNRDA 60

GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT FILING DATE: 1998-02-18
 PRIOR FILING DATE: 1998-02-18
 PRIOR FILING DATE: 1998-02-18
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 32242
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-991A-32242

Query Match 12.5%; Score 189; DB 2; Length 268;
 Best Local Similarity 25.5%; Pred. No. 5.3e-12;
 Matches 75; Conservative 38; Mismatches 115; Indels 66; Gaps 10;

11 LEVEEQHRRHPSMLLGLGQGLDWPPEFIRGLAERGRVPCFDRDAGLSTKLGK 70
 16 LNSLEGAGAPVLLSLGSLGDLGMDTQ-IPALTAH-FVRLYDTRGHGASLVTG-- 71

71 KPNARVLLASMGKLPVPYTLDDMALDTVGLMDALGIESHTVGVSMGMIAQILGAK 130
 72 -----PVAIGQAGVALLDLELPRVHFGLSGMLGQGLG 112

131 HGERVKSITLMTSS-----GNRPMAPAPQVL---QKFMVPSKMDKEWIKYNLEL 180
 113 AGERLGRVLNCAAKIAKIAKIAKIAKIAKIAKIAKIAKIAKIAKIAKIAKIAKIAK 162

181 LTTIGSPGLDREKALDVKSIERCLCEGTQQLAALQSGSRVKKLRRITAVTLVSG 240
 163 --TAGF--AEREPAQVRIVAMLAATSPOGVAANCAVRDADF-EQGLVQAPTLVAG 217

241 AEDPLPYQCCGRDIADHPIGAFELIEGMDHIDIPERHPIRIELTAGHAAAE 294
 218 SHDATTVPDARFQARIADA-----QLVEFAAAHLSNVEA 253

RESULT 12
 US-09-902-540-11639
 Sequence 11639, Application US/09902540
 Patent No. 6833447

GENERAL INFORMATION:
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Wiegand, Roger C.
 TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 FILE REFERENCE: 38-10(15849)B
 CURRENT FILING DATE: 2001-07-10
 PRIOR FILING DATE: 2000-07-10
 NUMBER OF SEQ ID NOS: 16825
 SEQ ID NO 11639
 LENGTH: 265
 TYPE: PRT
 ORGANISM: Myxococcus xanthus
 US-09-902-540-11639

Query Match 12.3%; Score 185; DB 2; Length 265;
 Best Local Similarity 26.7%; Pred. No. 1.4e-11;
 Matches 77; Conservative 38; Mismatches 121; Indels 52; Gaps 9;

1 MAQVANGITILEYEQGHRHPSMLLI--MGLGQGLDWPPEFIRGLAERGRVPCFDR 58
 1 MLTVTVGVPLHYRDEG--KGPPVLLHAFPLNGSTFD---KQVKAUSGR-YRFIIPDIR 54

59 DAGLSTKLEGVKKPNARVLLASMGKLPVPYTLDDMALDTVGLMDALGIESHTVGVCS 118
 55 GFGEALGDG-----PTMFRIARDALSLDLDALNDTIVVVGVS 93

119 MGGMIAQILGAKHGERVKSITLMT-----SSGNRPMAPAPQVLQKFMVPSKMDKEE 172
 94 MGYAAMALLREDAGRVSGLVLMTOATADDAEKKARRETSIAQALE-----VGVEP 145

173 WIKYNLELTTIGSPGLDREKALDVKSIERCLCEGTQQLAALQSGSRVKKLRRIA 232
 146 IQAMLPKMWAAAGPDSVPAQEV-----ALMRAASPAVAAALRGMALRPSKMDLARYA 200

233 VPTLVISGAEDPLPYQCCGRDIADHPIGAFELIEGMDHIDIPERHPIRIELTAGH 276
 201 GPALVIVGEHDALTTPPAKAKEIAGLISGAKLEVIPDAGHLANQEQOPOR 248

RESULT 13
 US-09-469-211A-6
 Sequence 6, Application US/09469211A
 Patent No. 6660524
 GENERAL INFORMATION:
 APPLICANT: J. Archer
 TITLE OF INVENTION: CONTROL OF GENE EXPRESSION IN EUKARYOTES
 FILE REFERENCE: 9341-021
 CURRENT FILING DATE: 1999-12-22
 PRIOR FILING DATE: 1999-12-22
 PRIOR FILING DATE: 1998-12-24
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: Patent in Ver. 2.1
 SEQ ID NO 6
 LENGTH: 289
 TYPE: PRT
 ORGANISM: Rhodococcus sp.
 US-09-469-211A-6

Query Match 12.1%; Score 183; DB 2; Length 289;
 Best Local Similarity 24.3%; Pred. No. 2.7e-11;
 Matches 68; Conservative 43; Mismatches 123; Indels 46; Gaps 9;

4 VKANGITILEYEQGHRHPSMLLGLGQGLDWPPEFIRGLAERG--FRVICDNRDAG 61
 22 IQAGPYRTRYLHAGDSSKPTLLILHGITH-----AEAYVRNLRSBHFVMAIDFIGHG 77

62 LSTKLEGVKKPNARVLLASMGKLPVPYTLDDMALDTVGLMDALGIESHTVGVSMGG 121
 78 YST-----KPDHPIEKHYIDHVLQLLDAIGVEKASFSGESLGG 116

122 MIAQILGAKHGERVKSITLMTSSGNRPMAPAPQVLQKFMVPSKMDKE---EWIKYNL 178
 117 WVTQAFAHDHPEKVDIVLN-TWGGT---MANPQVMERLYTSLSWEAKDPSWVRVKARL 171

179 ELL---TTIGSPGLDREKAL-----DVRKSIERCLCEGTQQLAALQSGSRVKKLRR 231
 172 EWLMDPTMTVTDLIRTRQAIFOQPDWLKACEMMALQDLETRKNMITDAT----LNGI 227

232 AVPTLVISGAEDPLPYQCCGRDIADHPIGAFELIEGMDH 271
 228 TVPAMVLWTTKDFSGPVDKRIASHIPGAKLAIMENCGH 267

RESULT 14
 US-09-446-681-6
 Sequence 6, Application US/09446681
 Patent No. 6849442
 GENERAL INFORMATION:
 APPLICANT: Archer, John AC
 APPLICANT: Summers, David K
 APPLICANT: Roland, Herve J
 APPLICANT: Powell, Justin AC
 TITLE OF INVENTION: Biosensor materials and methods

Query Match 12.3%; Score 185; DB 2; Length 265;
 Best Local Similarity 26.7%; Pred. No. 1.4e-11;
 Matches 77; Conservative 38; Mismatches 121; Indels 52; Gaps 9;

1 MAQVANGITILEYEQGHRHPSMLLI--MGLGQGLDWPPEFIRGLAERGRVPCFDR 58
 1 MLTVTVGVPLHYRDEG--KGPPVLLHAFPLNGSTFD---KQVKAUSGR-YRFIIPDIR 54

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 13, 2006, 13:10:48 ; Search time 161 Seconds
(without alignments)
762.993 Million cell updates/sec

Title: US-10-686-490D-2
Perfect score: 1509
Sequence: 1 MAQVKANGITLEYEEOGHRH.....ERHPLIELIAGHAAAAA 294

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA Main:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1509	100.0	294	5	US-10-686-490D-2
2	314.5	20.8	267	4	US-10-033-078-10
3	314.5	20.8	267	4	US-10-763-933-10
4	268.5	17.8	273	5	US-10-211-028-165
5	227.5	15.1	271	4	US-10-329-079-51
6	220.5	14.6	276	4	US-10-282-122A-49054
7	216.5	14.3	264	4	US-10-329-079-18
8	204	13.5	271	3	US-09-976-059-10
9	204	13.5	303	4	US-10-156-761-8055
10	198	13.1	232	3	US-09-896-578-4
11	198	13.1	232	3	US-09-971-490-14
12	198	13.1	232	4	US-10-289-148-4
13	198	13.1	232	4	US-10-038-854-66
14	198	13.1	232	4	US-10-193-452-36
15	198	13.1	232	4	US-10-193-452-87
16	197	13.1	217	4	US-10-051-874-95
17	193	12.8	224	4	US-10-099-322-303
18	193	12.8	224	4	US-10-044-564-303
19	191.5	12.7	272	5	US-10-211-028-8
20	189.5	12.6	251	4	US-10-068-134-4
21	185.5	12.3	233	4	US-10-056-744B-4
22	183.5	12.2	251	3	US-09-738-626-6126
23	183.5	12.2	251	5	US-10-494-675-50
24	183	12.1	289	5	US-10-732-859-6
25	181.5	12.0	264	4	US-10-724-972A-7349
26	180.5	12.0	258	4	US-10-282-122A-70509
27	180.5	12.0	310	4	US-10-425-115-292028

Sequence 78, Appl
Sequence 43203, A
Sequence 9245, Ap
Sequence 11944, A
Sequence 9087, Ap
Sequence 46, Appl
Sequence 45, Appl
Sequence 65472, A
Sequence 67314, A
Sequence 2, Appl
Sequence 70255, A
Sequence 8025, Ap
Sequence 66385, A
Sequence 181547, A
Sequence 9285, Ap
Sequence 266316, A
Sequence 54034, A
Sequence 23, Appl

28 175.5 11.6 289 4 US-10-354-437-78
29 175.5 11.6 339 4 US-10-282-122A-49203
30 174.5 11.6 379 4 US-10-156-761-9245
31 171.5 11.4 261 4 US-10-156-761-11944
32 171 11.3 266 4 US-10-156-761-9087
33 171 11.3 7349 4 US-10-314-657-46
34 171 11.3 7349 5 US-10-473-193-46
35 168.5 11.2 249 4 US-10-282-122A-65472
36 168 11.1 430 5 US-10-362-530-2
37 168 11.1 510 5 US-10-362-530-2
38 167 11.1 262 4 US-10-282-122A-70255
39 167 11.1 575 5 US-10-739-930-8025
40 164.5 10.9 491 4 US-10-425-114-66385
41 163.5 10.8 639 4 US-10-437-963-181547
42 163 10.8 347 4 US-10-156-761-9285
43 160.5 10.6 477 4 US-10-425-115-266316
44 160.5 10.6 522 4 US-10-425-114-54034
45 159 10.5 299 3 US-09-942-025-23

ALIGNMENTS

RESULT 1
US-10-686-490D-2
; Sequence 2, Application US/10686490D
; Publication No. US20050153404A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Anti-Kazlauskas-Lipases
; FILE REFERENCE: Lea 35 991
; CURRENT APPLICATION NUMBER: US/10/686,490D
; CURRENT FILING DATE: 2003-10-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 2
; LENGTH: 294
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: bacterial
US-10-686-490D-2

Query Match 100.0%; Score 1509; DB 5; Length 294;
Best Local Similarity 100.0%; Pred. No. 3.6e-145;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQVKANGITLEYEEOGHRHPSMLLMGGQLIDWPEEFIRGLAERGFVCFDNRDA 60
DB 1 MAQVKANGITLEYEEOGHRHPSMLLMGGQLIDWPEEFIRGLAERGFVCFDNRDA 60
QY 61 GLSTKLEGVKKNIA RVFLASMLKPRVPYPTLDDMALDTVGLMDALGISTHVGVSMG 120
DB 61 GLSTKLEGVKKNIA RVFLASMLKPRVPYPTLDDMALDTVGLMDALGISTHVGVSMG 120
QY 121 GMAIQLGAKHGSRVKS LTIMTSSGNRPMAPRPOVLQKFMVPKSMDEEIKYNLEL 180
DB 121 GMAIQLGAKHGSRVKS LTIMTSSGNRPMAPRPOVLQKFMVPKSMDEEIKYNLEL 180
QY 181 LTTIGSPGLDREKALDVKSIERCICPEGTORQLAAILOSGRSVKLLRIAPVTLVISG 240
DB 181 LTTIGSPGLDREKALDVKSIERCICPEGTORQLAAILOSGRSVKLLRIAPVTLVISG 240
QY 241 AEDPLLPYOCGRDIADHIFGARFELIEGMDHIDPERHIFRLIELIAGHAAAAA 294
DB 241 AEDPLLPYOCGRDIADHIFGARFELIEGMDHIDPERHIFRLIELIAGHAAAAA 294

RESULT 2
US-10-033-078-10
; Sequence 10, Application US/10033078
; Publication No. US20020123109A1
; GENERAL INFORMATION:

APPLICANT: Furuichi, Yasuhiro
APPLICANT: Hoshino, Tatsuo
APPLICANT: Kimura, Hitoshi
APPLICANT: Kiyasu, Tatsuya
APPLICANT: Nagahashi, Yoshie
TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
FILE REFERENCE: Biotin Genes
CURRENT APPLICATION NUMBER: US/10/033,078
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 09/594,185
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: EP 96115540.5
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 267
TYPE: PRT
ORGANISM: Kurthia sp.
US-10-033-078-10

Query Match 20.8%; Score 314.5; DB 4; Length 267;
Best Local Similarity 33.1%; Pred. No. 2.9e-23;
Matches 92; Conservative 43; Mismatches 104; Indels 39; Gaps 9;
QY 1 MAQVKANGITLEYEEQGRHHPMSLLIMGLGQLIDWPPEFIRGLAERGFVCFDNRDA 60
DB 1 MPFVNDNESLYYEV--HGQGDPLLLIMGLGYNLSW-HRTVPTLAKR-FKVIIVFDNRGV 56
QY 61 GLSTKLEGVKKPNRIARVFLASMLGLKPRVPTLDDMALDVTGLMDALGIESTHVGVSMG 120
DB 57 GKSS-----KPEQPYSIEMMAEDARAVLDAVSDSAHVYGISMG 95
QY 121 GMIAQILGAKHGERVKSILMITSSGNPRMPAPRPQV---LQKFMVPKSMDEEW--- 173
DB 96 GMIAQRLAITYPERVRSVLGCTTAGTTHIQPSPEISILMVRASITGSPRDNAMLAAP 155
QY 174 IKYNLELLTTIGSPGLDREKLALDVKRSIERCLCPEGTQRLAAILOSGSRVKLLRRIAV 233
DB 156 IVYSQAFIEK--HPELIQE---DIQRIEIIPTPSAYLSQLQACLTHDTSNE-LDKINI 208
QY 234 PTLVISGAEDPLLPYOCGRDIADHIPGARFELIEGMGH 271
DB 209 PTLIHGDADNLVPYENGKMLAERIQGSQFHTVSCAGH 246

RESULT 3
US-10-763-933-10
Sequence 10, Application US/10763933
Publication No. US20040137584A1
GENERAL INFORMATION:
APPLICANT: Furuichi, Yasuhiro
APPLICANT: Hoshino, Tatsuo
APPLICANT: Kimura, Hitoshi
APPLICANT: Kiyasu, Tatsuya
APPLICANT: Nagahashi, Yoshie
TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
FILE REFERENCE: Biotin Genes
CURRENT APPLICATION NUMBER: US/10/763,933
CURRENT FILING DATE: 2004-01-23
PRIOR APPLICATION NUMBER: US/10/033,078
PRIOR FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 09/594,185
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: EP 96115540.5
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 267
TYPE: PRT
ORGANISM: Kurthia sp.
US-10-763-933-10

Query Match 20.8%; Score 314.5; DB 4; Length 267;
Best Local Similarity 33.1%; Pred. No. 2.9e-23;
Matches 92; Conservative 43; Mismatches 104; Indels 39; Gaps 9;
QY 1 MAQVKANGITLEYEEQGRHHPMSLLIMGLGQLIDWPPEFIRGLAERGFVCFDNRDA 60
DB 1 MPFVNDNESLYYEV--HGQGDPLLLIMGLGYNLSW-HRTVPTLAKR-FKVIIVFDNRGV 56
QY 61 GLSTKLEGVKKPNRIARVFLASMLGLKPRVPTLDDMALDVTGLMDALGIESTHVGVSMG 120
DB 57 GKSS-----KPEQPYSIEMMAEDARAVLDAVSDSAHVYGISMG 95
QY 121 GMIAQILGAKHGERVKSILMITSSGNPRMPAPRPQV---LQKFMVPKSMDEEW--- 173
DB 96 GMIAQRLAITYPERVRSVLGCTTAGTTHIQPSPEISILMVRASITGSPRDNAMLAAP 155
QY 174 IKYNLELLTTIGSPGLDREKLALDVKRSIERCLCPEGTQRLAAILOSGSRVKLLRRIAV 233
DB 156 IVYSQAFIEK--HPELIQE---DIQRIEIIPTPSAYLSQLQACLTHDTSNE-LDKINI 208
QY 234 PTLVISGAEDPLLPYOCGRDIADHIPGARFELIEGMGH 271
DB 209 PTLIHGDADNLVPYENGKMLAERIQGSQFHTVSCAGH 246

RESULT 4
US-10-211-028-165
Sequence 165, Application US/10211028
Publication No. US20050027113A1
GENERAL INFORMATION:
APPLICANT: CUBIST PHARMACEUTICALS, INC.
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO THE DAPTOMYCIN
FILE OF INVENTION: BIOSYNTHETIC GENE CLUSTER
FILE REFERENCE: CUB-12 PCT CIP
CURRENT APPLICATION NUMBER: US/10/211,028
CURRENT FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: PCT/US02/24310
PRIOR FILING DATE: 2002-10-25
PRIOR APPLICATION NUMBER: PCT/US01/32354
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 60/310,385
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 60/379,866
PRIOR FILING DATE: 2002-05-10
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 165
LENGTH: 273
TYPE: PRT
ORGANISM: Streptomyces coelicolor
US-10-211-028-165

Query Match 17.8%; Score 268.5; DB 5; Length 273;
Best Local Similarity 26.8%; Pred. No. 1.5e-18;
Matches 85; Conservative 44; Mismatches 109; Indels 79; Gaps 10;
QY 1 MAQVKANGITLEYEEQ---GHRHHPMSLLIMGLGQLIDWPPEFIRGLAERGFVCFD 56
DB 1 MPVLTGVNGIRINYDDAPPAGQANAPAVLLVNGSGSGRAWHLHQVPALVAAGFRVISFD 60
QY 57 NRDAGLSTKLEGVKKPNRIARVFLASMLGLKPRVPTLDDMALDVTGLMDALGIESTHVVG 116
DB 61 NRGIAPESECPG-----GFGIDDLVADTAALVEELRIGPCRVAG 99
QY 117 VSWGGMIAQILGAKHGERVKSILMITSSGNP-----RMPAPRPQV 157
DB 100 ISMGAHIAQELALSRLPDLVRLVLMATRA-REDALREALCRAEMELYDOGIRLPAAVEAV 158
QY 158 LQKFMV-PKSMDEK-----EWIKYNLELLTTIGSPGLDREKLALDVKRSIERCLCPEGTQ 212
DB 159 VOAMQNLSRPTLDNDVQARDWLDV-LELTRRSGA-----GTR 194

QY 213 KOLAALQSGSRVKKLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELIEGMGH- 271
 Dp 195 AQL-GVRVDGRREAYGIRATRVAFQDDLIAPPHLGREVAIPAEGAYELVPDCGHY 253
 QY 272 ---DIPERHIFRLIELI 285
 Db 254 GYLESPPDAVNSLVEFL 270

RESULT 5
 US-10-329-079-51
 ; Sequence 51, Application US/10329079
 ; Publication No. US20030198981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FARNET, Chris
 ; APPLICANT: ZAZOPOULOS, Emmanuel
 ; APPLICANT: STAFFA, Alfredo
 ; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF LIPOPEPTIDES
 ; CURRENT APPLICATION NUMBER: US/10/329,079
 ; CURRENT FILING DATE: 2002-12-24
 ; NUMBER OF SEQ ID NOS: 66
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 51
 ; LENGTH: 271
 ; TYPE: PRT
 ; ORGANISM: Streptomyces refuineus
 US-10-329-079-51

Query Match 15.1%; Score 227.5; DB 4; Length 271;
 Best Local Similarity 28.9%; Pred. No. 2.2e-14;
 Matches 84; Conservative 47; Mismatches 98; Indels 83; Gaps 13;

QY 1 MAQVKANGITLEYEEQGRHHPHSMLLIMGLGQLIDWPPEEFIRGLAERGFVCFDNRDA 60
 Db 1 MPTRINGIALDHRTG--SGPPVLLINGSAAKSAHLLHQVPAVAGFEAVTFNR-- 56
 QY 61 GLSTKLGKVPKPNIAVFLVLLASMGKPR---VPVTLDDMALDVTGLMDALGISTHVVGV 117
 Db 57 -----GVPPSGGGPGFTGLGMAADVTGLTLEHLGIGPCAVVGM 93
 QY 118 SMGGMIAQILCAKGERVKSLTMITSSGNRMPA-----PRPVQLQK 160
 Db 94 SLGARVAREARTPDILVSRCLVAPRARSDRMPAACTAAEIALADSGVTLPPRYRAVVR 153
 QY 161 FMR--VPKSM--DKX--EWIKYNLELLTTIGSPGLDREKLALDVKRSIERCLCEGTORQ 214
 Db 154 AMQNLSPRTLADRDQADWLDV-LELAADG-FGL-----RTQLE 191
 QY 215 LAALQSGSRVKKLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELIEGMGH--- 271
 Db 192 LSA---ADDRGEDLAGITAPCRVIAFADDIVAPPHLAKETADALPEADYHVVPDCGHY 248
 QY 272 -DIPERHIFRLI 282
 Db 249 LERPDR-VNRLI 259

RESULT 6
 US-10-282-122A-49054
 ; Sequence 49054, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 49054
 ; LENGTH: 276
 ; TYPE: PRT
 ; ORGANISM: Burkholderia fungorum
 US-10-282-122A-49054

Query Match 14.6%; Score 220.5; DB 4; Length 276;
 Best Local Similarity 28.8%; Pred. No. 1.2e-13;
 Matches 84; Conservative 43; Mismatches 104; Indels 63; Gaps 12;

QY 3 QVKANGITLEYEEQGRHHPHSMLLIMGLGQLIDWPPEEFIRGLAERGFVCFDNRDAGL 62
 Db 8 ETTVRGQTSYLELGDGPR-TLLLLHGITSASNWLST-MPALAQGRVRIAPDLQGFQ 65
 QY 63 STKLGKVPKPNIAVFLVLLASMGKPRVPVTLDDMALDVTGLMDALGISTHVVGVSMGM 122
 Db 66 SS-----KPSV-----PVRPRTLSDM---VAPLLDALGLEMVSIVGQSMGGH 104
 QY 123 IAQILCAKGERVKSLTMITSSG--NPRMPAPR-----PQVLOKEMRVPKSMCKE-- 171
 Db 105 VAGLFAAQYPERVEALVYNAGVGLALPEVKDPRDLGHAVTPGGL--WALNPATRDDSR 162
 QY 172 --EWIKYNLELLTTIGSPGLDREKLALDVKRSIERCLCEGTORQLAAILQSGSR----- 224
 Db 163 LLEWVFHDQGLVTEELIDGFYADRLG-----KGDGAVIRSISESWARREDTL 209
 QY 225 ----VKLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELIEGMGHDP 274
 Db 210 ESAFTGLERR---PVLVIOARQDKVAPYHLGRAIHEGIAGSREVLEDCGHAPP 260

RESULT 7
 US-10-329-079-18
 ; Sequence 18, Application US/10329079
 ; Publication No. US20030198981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FARNET, Chris
 ; APPLICANT: ZAZOPOULOS, Emmanuel
 ; APPLICANT: STAFFA, Alfredo
 ; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF LIPOPEPTIDES
 ; FILE REFERENCE: 3002-11US
 ; CURRENT APPLICATION NUMBER: US/10/329,079
 ; CURRENT FILING DATE: 2002-12-24
 ; NUMBER OF SEQ ID NOS: 66

```

Qy 124 AQLGAKHGERVKSLLTMITSSGNPRMPAPRPQ-----VLQFMRVPKSMDK-----170
Db 106 AQLAVDHPHVLRCVLIATLAPDFAARAQNOADIDLLSEGVTLPAAYEAATAVFKWFS 165
Qy 171 -----EEWIKYNLELLTTIGSPGLDREKLALDVRKSIERCLCPECTQRLAAILQ 220
Db 166 PATLNDVAVREWDI-FELSGTGVSGAG-----GQWAAEL-- 199
Qy 221 SGRVSKLLRRIAPVTLTIVISGAEDPLLPYOCGRDIADHIFGARFELIEGMGH-----DIPER 276
Db 200 TGDRAALRSVTAFCRVISFADDDLIITPHLAEVAEAIPOCDLVEISRCGHGLGVLERPDA 259
Qy 277 HIPRLIELIAGH 288
Db 260 VNAAILEFLDSH 271

RESULT 9
US-10-156-761-8055
; Sequence 8055, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761*
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8055
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8055

Query Match 13.5%; Score 204; DB 4; Length 303;
Best Local Similarity 27.1%; Pred. No. 6.6e-12;
Matches 76; Conservative 46; Mismatches 102; Indels 56; Gaps 10;

Qy 22 PSMLLIMLGGQLIDWPEEFIRGLAERGPVCFDNRDAGLSLKLGKGVKKPNRIARVFLLA 81
Db 65 PALIFHYVGGSSRTWI PVQLRDPGQGF--VAYDQRHGGSTSVFG-----109
Qy 82 SMGLKRPVPTTLDNALDTVGLMDALGIESTHVGVSMGGMIAQILGAKHGERVKSLLTLM 141
Db 110 -----PYDLEQLADDAQRVVDALGYSRYVLVGHSMGKQVAKIILAAKRPAGLGVVLV 161
Qy 142 ITSSGNPRPAPRPQVLQKFMVPKSMKKEWIKYNLELLTTIG--SPGLDREKLALDVR 199
Db 162 A-----PAPFAPIGVTEQVQVTSYHADNNEAVLQSIDLMLTRGGUTPELRQVVEDSLR 216
Qy 200 KSIERCLCPEGTQRLAAILQSGSRVKLLRRIAPVTLTIVISGAEDPLLPYOCGRDIADH-- 257
Db 217 GGDEARL--EWPFRGLVQDVSG-----VSAIEVPVLVLAGSHDKVDPPPTV---LADHLL 266
Qy 258 --IPGARFELIEGMGH-----DIPERHIPRLIELIAGHAAA 291
Db 267 PLIPTATLTVLKDTHLSPLEVPDQ-----VAAHIGA 298

RESULT 10
US-09-896-578-4
Sequence 4, Application US/09896578

```



```
Patent No. US20020077310A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 32225, A NOVEL HUMAN ALPHA BETA HYDROLASE
; TITLE OF INVENTION: FAMILY MEMBER
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 10448-073001
; CURRENT APPLICATION NUMBER: US/09/896,578
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/214,948
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-896-578-4

Query Match      13.1%; Score 198; DB 3; Length 232;
Best Local Similarity 31.5%; Pred. No. 1.9e-11;
Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;

Qy 50 FRVICFNRDAGLSTKLEGVKKNIAKVFLVLLASMLGKPRVPYTLDDMALDTVGLMDALGI 109
Db 1 FRVILLDLRGFGESSPSD-----LAE-----YRFDDLAEDLEALLDALGL 40
Qy 110 ESTHV-VGVSMGGMIAQILGAKHGE-RVKSLLTMITSSGNPRMPA-----PRPQVLQK 160
Db 41 EKPVLVGHSMGGAIALAYAAKYPELRVKALVLV-----SPPLPAGLSSDLFFRQGNLEG 95
Qy 161 FM-----RVPKSMDEXEIKYNLELLTTIGSPGL-DREKALDVRKSIERCLCPECTOR 213
Db 96 LLANFNRLSRV--EALLGRALKQFFLLGRPLVSDFLKQAEWLSLIRQGEDDGDG 153
Qy 214 QLAAILQSGSRVK-----LLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELIEGM 269
Db 154 LLGNAVALGKLLQWDLKALDKIKVPTLVITWGTDDPLVPLDASEKLSALIPNAEVVVIDDA 213
Qy 270 GH-----DIPERHIPRLIELI 285
Db 214 GHLALLEKPE-EVAELIKFL 232

RESULT 12
US-10-289-148-4
; Sequence 4, Application US/10289148
; Publication No. US2003010001A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Spurling, Heidi L.
; TITLE OF INVENTION: 46694, A HUMAN ALPHA/BETA HYDROLASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREFOR
; FILE REFERENCE: MP101-295PIRNM
; CURRENT APPLICATION NUMBER: US/10/289,148
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: US 60/334,225
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Alpha/Beta Hydrolase Fold Consensus Domain
US-10-289-148-4

Query Match      13.1%; Score 198; DB 4; Length 232;
Best Local Similarity 31.5%; Pred. No. 1.9e-11;
Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;

Qy 50 FRVICFNRDAGLSTKLEGVKKNIAKVFLVLLASMLGKPRVPYTLDDMALDTVGLMDALGI 109
Db 1 FRVILLDLRGFGESSPSD-----LAE-----YRFDDLAEDLEALLDALGL 40
Qy 110 ESTHV-VGVSMGGMIAQILGAKHGE-RVKSLLTMITSSGNPRMPA-----PRPQVLQK 160
Db 41 EKPVLVGHSMGGAIALAYAAKYPELRVKALVLV-----SPPLPAGLSSDLFFRQGNLEG 95
Qy 161 FM-----RVPKSMDEXEIKYNLELLTTIGSPGL-DREKALDVRKSIERCLCPECTOR 213
Db 96 LLANFNRLSRV--EALLGRALKQFFLLGRPLVSDFLKQAEWLSLIRQGEDDGDG 153
Qy 214 QLAAILQSGSRVK-----LLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELIEGM 269
Db 154 LLGNAVALGKLLQWDLKALDKIKVPTLVITWGTDDPLVPLDASEKLSALIPNAEVVVIDDA 213
Qy 270 GH-----DIPERHIPRLIELI 285
Db 214 GHLALLEKPE-EVAELIKFL 232

Patent No. US20020077310A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 32225, A NOVEL HUMAN ALPHA BETA HYDROLASE
; TITLE OF INVENTION: FAMILY MEMBER
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 10448-073001
; CURRENT APPLICATION NUMBER: US/09/896,578
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/214,948
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-896-578-4

Query Match      13.1%; Score 198; DB 3; Length 232;
Best Local Similarity 31.5%; Pred. No. 1.9e-11;
Matches 82; Conservative 36; Mismatches 90; Indels 52; Gaps 12;

Qy 50 FRVICFNRDAGLSTKLEGVKKNIAKVFLVLLASMLGKPRVPYTLDDMALDTVGLMDALGI 109
Db 1 FRVILLDLRGFGESSPSD-----LAE-----YRFDDLAEDLEALLDALGL 40
Qy 110 ESTHV-VGVSMGGMIAQILGAKHGE-RVKSLLTMITSSGNPRMPA-----PRPQVLQK 160
Db 41 EKPVLVGHSMGGAIALAYAAKYPELRVKALVLV-----SPPLPAGLSSDLFFRQGNLEG 95
Qy 161 FM-----RVPKSMDEXEIKYNLELLTTIGSPGL-DREKALDVRKSIERCLCPECTOR 213
Db 96 LLANFNRLSRV--EALLGRALKQFFLLGRPLVSDFLKQAEWLSLIRQGEDDGDG 153
Qy 214 QLAAILQSGSRVK-----LLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARFELIEGM 269
Db 154 LLGNAVALGKLLQWDLKALDKIKVPTLVITWGTDDPLVPLDASEKLSALIPNAEVVVIDDA 213
Qy 270 GH-----DIPERHIPRLIELI 285
Db 214 GHLALLEKPE-EVAELIKFL 232

RESULT 11
US-09-971-490-14
; Sequence 14, Application US/09971490
; Publication No. US20030032091A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 48120, 23479 AND 46689, NOVEL HUMAN HYDROLASES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 10448-100001
; CURRENT APPLICATION NUMBER: US/09/971,490
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,170
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/237,991
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
US-09-971-490-14
```



```

; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/267,054
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: US 09/888,911
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: PCT/US01/19967
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,688
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-193-452-87

```

```

Query Match      13.1%; Score 198; DB 4; Length 232;
Best Local Similarity 31.5%; Pred. No. 1.9e-11;
Matches 82; Conservative 36; Mismatches 50; Indels 52; Gaps 12;

Qy      50 FRVICFNRDAGLSTKLEGVKKNPNIARVFLASMGKLPKRVPTLDDMALDTVGLMDALGI 109
Db      1 FRVILLDLRGFGESSPSD-----LAE-----YRFDLAEDELEALLDALGL 40

Qy      110 ESTHV-VGVSMGMIQIILGAKHGE-RVKSLLTMITSSGNPRMPA-----PRPQVLQK 160
Db      41 EKPVLVGHSMGGAIALAYAAKYPRLRYKALVLV-----SPPLPAGLSSDLFPRQGNLEG 95

Qy      161 FM-----RVPKSMKKEWIKYNLELLTTIGSPGL-DREKLALDVKXSIERCLCEGTQR 213
Db      96 LLLANPRNLSRSV--EALLGRALKQFLLGRPLVSDFLKQAEWLSLIRQGEDDGDG 153

Qy      214 QLAAILQSGSRVK----LLRRIAYPTLVISGAEDPLLPYQCGRDIADHIGARFELIEGM 269
Db      154 LLGAVALGKLLQWDLKALDKIVPTLVINGTDDPLVPLDASEKLSALIPNAEVVVWIDDA 213

Qy      270 GH-----DIPERHPIRLIELI 285
Db      214 GHLELEKPE-EVAELIKFL 232

```

Search completed: April 13, 2006, 13:14:23
Job time : 162 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Ran on: April 13, 2006, 13:11:49 ; Search time 27 Seconds
(without alignments)
462.661 Million cell updates/sec

Title: US-10-686-490D-2
Perfect score: 1509
Sequence: 1 MAQVXANGITLVEEEOGHRH.....ERHPRILIELIAGHAAAEA 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SIDSS5/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /SIDSS5/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 3: /SIDSS5/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 4: /SIDSS5/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 5: /SIDSS5/ptodata/1/pubpaa/US03_NEW_PUB pep.*
- 6: /SIDSS5/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 7: /SIDSS5/ptodata/1/pubpaa/US11_NEW_PUB pep.*
- 8: /SIDSS5/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	209.5	13.9	260	7	US-11-087-099-1714
2	204	13.5	271	7	US-11-205-109-10
3	201.5	13.4	312	7	US-11-087-099-5676
4	193	12.8	224	7	US-11-054-281-303
5	187	12.4	265	7	US-11-087-099-4437
6	183.5	12.2	251	7	US-11-055-822-470
7	183	12.1	265	7	US-11-087-099-9057
8	176.5	11.7	266	7	US-11-087-099-967
9	176.5	11.2	266	7	US-11-087-099-8501
10	168.5	11.2	293	6	US-10-467-657-6664
11	160	10.6	2197	7	US-11-075-185-8
12	155.5	10.3	386	7	US-11-179-977-9
13	147	9.7	308	7	US-11-096-568A-25601
14	147	9.7	330	7	US-11-096-568A-25600
15	145.5	9.6	264	7	US-11-096-568A-29097
16	145.5	9.6	307	7	US-11-096-568A-29096
17	145.5	9.6	311	7	US-11-096-568A-29095
18	141.5	9.4	256	7	US-11-096-568A-29086
19	141.5	9.4	267	7	US-11-087-099-11146
20	141.5	9.4	299	7	US-11-096-568A-29085
21	141.5	9.4	303	7	US-11-096-568A-29084
22	136.5	9.0	269	7	US-11-179-977-11
23	136.5	9.0	289	7	US-11-087-099-8151
24	135	8.9	311	7	US-11-096-568A-4794
25	130.5	8.6	256	7	US-11-179-977-15

26	128	8.5	250	7	US-11-096-568A-33303	Sequence 33303, A
27	128	8.5	273	7	US-11-096-568A-33302	Sequence 33302, A
28	128	8.5	282	7	US-11-096-568A-33301	Sequence 33301, A
29	124	8.2	267	7	US-11-096-568A-4795	Sequence 4795, Ap
30	122	8.1	331	6	US-10-454-437-226	Sequence 226, App
31	121.5	8.1	302	7	US-11-087-099-10293	Sequence 10293, A
32	120.5	8.0	311	6	US-10-467-657-5644	Sequence 5644, Ap
33	120.5	8.0	334	6	US-10-525-674-30	Sequence 30, Appl
34	119	7.9	293	7	US-11-096-568A-13018	Sequence 13018, A
35	117.5	7.8	487	7	US-11-096-568A-13017	Sequence 13017, A
36	117.5	7.8	546	7	US-11-096-568A-13017	Sequence 13017, A
37	117	7.8	284	7	US-11-096-568A-5989	Sequence 5989, Ap
38	117	7.8	401	7	US-11-096-568A-5988	Sequence 5988, Ap
39	117	7.8	416	7	US-11-096-568A-5987	Sequence 5987, Ap
40	114.5	7.6	311	7	US-11-096-568A-6768	Sequence 6768, Ap
41	113	7.5	324	7	US-11-096-568A-33701	Sequence 33701, A
42	113	7.5	380	6	US-10-525-674-28	Sequence 28, Appl
43	111.5	7.4	259	7	US-11-087-099-6077	Sequence 6077, Ap
44	110.5	7.3	269	7	US-11-087-099-9507	Sequence 9507, Ap
45	110.5	7.3	399	7	US-11-096-568A-21969	Sequence 21969, A

ALIGNMENTS

RESULT 1
US-11-087-099-1714
; Sequence 1714, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1714
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Bradyrhizobium japonicum USDA 110
US-11-087-099-1714

Query Match 13.9%; Score 209.5; DB 7; Length 260;
Best Local Similarity 26.1%; Pred. No. 4.1e-12;
Matches 74; Conservative 37; Mismatches 109; Indels 63; Gaps 8;

Qy	1	MAQVXANGITLVEEEOGHRHPSMLLMGGLIDMPPEFIRGLAERGFVFCFNRDA	60
Db	1	MEMIDADGCLINVSVEGRDGGPTLMLNSLGGCTLQMW-EPQMKALTO-VFRVIRYDRGH	58
Qy	61	GLSTKLEGVKKPNIAVFLASMGKLPVYPYTLDDMALDTVGLMDALGIESTHVGVSMG	120
Db	59	G-----KSNV-----PPGPTWMEFRDVLAILNDLNIKVKHVGCLSMG	97
Qy	121	GMAIQLKAGHGVKSLTLMITSSGNPNPAPRPOVLQKFMKVPKSMKEIKYNLEL	180
Db	98	GMVGWGLGANAPERFGKLIANTSC-----YYAETPKW-----LER	133
Qy	181	LTTIGSPGL-----DREKLADVRKSTERGLCPGEGTORQAAA--ILQSGSR	224
Db	134	IDAVKKGGAADVAVIAGWLQDFRERPDITARMKAMLLASPVGYLACCALSTLDQ	193
Qy	225	VKLLRRIAVPTLVISGAEDPLLPYQCGRDIADHIFGARGFELIE	267
Db	194	RELLPKIKPTLVIAGRHDMATPISAGELIRSKIFGANNTIID	236

RESULT 2
US-11-205-109-10
; Sequence 10, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Parnet, Chris

APPLICANT: Zazopoulos, Emmanuel
 APPLICANT: Staffa, Alfredo
 TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
 FILE REFERENCE: 3002-ZUS
 CURRENT APPLICATION NUMBER: US/11/205,109
 CURRENT FILING DATE: 2005-08-17
 PRIOR APPLICATION NUMBER: US/09/976,059
 PRIOR FILING DATE: 2001-10-15
 PRIOR APPLICATION NUMBER: US 60/239,924
 PRIOR FILING DATE: 2000-10-13
 NUMBER OF SEQ ID NOS: 46
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 10
 LENGTH: 271
 TYPE: PRT
 ORGANISM: Actinoplanes sp.
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1)-(1)
 OTHER INFORMATION: V represents a non-standard initiator codon. It is expected that the biosynthesized protein will have a formylmethionine residue
 OTHER INFORMATION: at this position
 US-11-205-109-10

Query Match 13.5%; Score 204; DB 7; Length 271;
 Best Local Similarity 25.0%; Pred. No. 1.4e-11; Mismatches 120; Indels 76; Gaps 9;
 Matches 78; Conservative 38;

QY 4 VKANGITLEYEEOGHRHPSMLLMGLGGQLIDWPPEFIRGLAERGRVFCFNRDAGLS 63
 DB 9 VTTNGVRLAYRRSGAGE--PVLIMMGSSAGQVTWHTQTPALHEAGYSTVVFDSR----- 61
 QY 64 TKLEGVKPNRIARVFLASMGKLPVPYTTLDMDALDTVGLMDALGIESTHVGVSMGMI 123
 DB 62 -----GIPPSDV-----PAGKYSLADMTADTGLIEALDLAPCRIVGTSLGAMI 105
 QY 124 AQILGAKHGERVKSLTMITSSGNPRMPAPRPQ-----VLQKFMVPKSMDK----- 170
 DB 106 AQELAVDPHELVRCAVLIATLRPDAAQAQNOADIDLLSGVTLPAAAYEAATAVFKMFS 165
 QY 171 -----EEWIKYNLELTITGSPGLDREKALDVRSIERCLCPEGTQRQLAAILQ 220
 DB 166 PATLNDVDVAVREWLDI-FELSGTGSAG-----GQAWAEL- 199
 QY 221 SGRSVKLLRRIAVPTLVISGAEDPLLPYQCGRIADHIFGAFELIEGMGH-----DIPER 276
 DB 200 TGDRAALRSVTPACRVISFADLLITPPHLLAAEVAEAIPDCDLVEISRCGHGLYLERPDA 259
 QY 277 HIPRLIELIAGH 288
 DB 260 VNAAILFELSH 271

RESULT 3
 US-11-087-099-5676
 Sequence 5676, Application US/11087099
 Publication No. US20060041961A1
 GENERAL INFORMATION:
 APPLICANT: Abad, Mark S. et al.
 TITLE OF INVENTION: Genes and Uses for Plant Improvement
 FILE REFERENCE: 38-21(53450)B EP
 CURRENT APPLICATION NUMBER: US/11/087,099
 CURRENT FILING DATE: 2005-03-22
 NUMBER OF SEQ ID NOS: 12464
 SEQ ID NO 5676
 LENGTH: 312
 TYPE: PRT
 ORGANISM: Rhodobacter sphaeroides
 US-11-087-099-5676

Query Match 13.4%; Score 201.5; DB 7; Length 312;
 Best Local Similarity 24.4%; Pred. No. 3e-11;
 Matches 71; Conservative 42; Mismatches 105; Indels 73; Gaps 8;

QY 1 MAQVKANGITLEYEEOGHRHPSMLLMGLGGQLIDWPPEFIRGLAERGRVFCFNRDA 60
 DB 51 MPHIEIGDURLNYLDEGRATGAPVVFHAHUGSLRIW--DNLIPLPDQDLRLVRLURGH 108
 QY 61 GLSTKLEGVKPNRIARVFLASMGKLPVPYTTLDMDALDTVGLMDALGIESTHVGVSMG 120
 DB 109 GRST-----TPEPPYANGALIRDAERLMEALSIREAVFVGCISIG 147
 QY 121 GMTAQILGAKHGERVKSLTMITSSGNPRMPAPRPQVLQKFMVPKSMDKEWIK-YNLE 179
 DB 148 GMTAQGLAVKRLDLVRGLVLCDTAA-----KIGTPEIWQ-----DRIDQVRSYGLE 193
 QY 180 LTTIGSPGLDREKALDVRSIERCLCP---EGTQRQL-----RAILQ 220
 DB 194 SLAD-----PTMKRWPAFAFROGPEGQLWRRFIEGDDPGYAGGAAIA 237
 QY 221 SGRSVKLLRRIAVPTLVISGAEDPLLPYQCGRIADHIFGAFELIEGMGH 271
 DB 238 GTDFYTTTARLTPTLALVSGEDGSTPPDLVRETASLIRGSRFEIVRGAGH 288

RESULT 4

US-11-054-281-303
 Sequence 303, Application US/11054281
 Publication No. US20060013813A1
 GENERAL INFORMATION:
 APPLICANT: Mezes et al.
 TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 FILE REFERENCE: 21402-240CIP
 CURRENT APPLICATION NUMBER: US/11/054,281
 CURRENT FILING DATE: 2005-02-08
 PRIOR APPLICATION NUMBER: 60/261,014
 PRIOR FILING DATE: 2001-01-11
 PRIOR APPLICATION NUMBER: 60/261,018
 PRIOR FILING DATE: 2001-01-11
 PRIOR APPLICATION NUMBER: 60/318,410
 PRIOR FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: 60/261,013
 PRIOR FILING DATE: 2001-01-11
 PRIOR APPLICATION NUMBER: 60/261,026
 PRIOR FILING DATE: 2001-01-11
 PRIOR APPLICATION NUMBER: 60/261,029
 PRIOR FILING DATE: 2001-01-11
 PRIOR APPLICATION NUMBER: 60/313,170
 PRIOR FILING DATE: 2001-08-17
 PRIOR APPLICATION NUMBER: 10/044,564
 PRIOR FILING DATE: 2002-01-11
 NUMBER OF SEQ ID NOS: 324
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 303
 LENGTH: 224
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: consensus
 OTHER INFORMATION: sequence
 US-11-054-281-303

Query Match 12.8%; Score 193; DB 7; Length 224;
 Best Local Similarity 28.1%; Pred. No. 1.2e-10;
 Matches 71; Conservative 31; Mismatches 71; Indels 80; Gaps 9;

QY 52 VICFDRDAGLSTKLEGVKPNRIARVFLASMGKLPVPYTTLDMDALDTVGLMDALGIES 111
 DB 1 VILFDLRGFGQSSPSD-----LAE-----YRPDDLAEDELALDLGLDK 40
 QY 112 THVGVSMGGMIAQILGAKHGERVKSLTMITSSGNPRMPAPRPQVLQKFMVPKSM--- 168
 DB 41 VILVGHSMGGAIAAAYAAKTPERKAVL-----VSAPHPALLSSRL-FPRNLFG 90
 QY 169 -----DKEEWIKYNLELTITGSPGLDREKALDVRSIERCLCPEGTQRQLAA 217

